AD-A034 521 ARMY CONCEPTS ANALYSIS AGENCY BETHESDA MD F/G 12/2
AN APPLICATION OF THE MULTIVARIATE ANALYSIS SYSTEM TO TAGS.(U)
DEC 76 D A NUSSBAUM
UNCLASSIFIED CAA-TP-76-3 NL

1 OF 1
AD
A034521

END
DATE
FILMED
2-77

DISCLAIMER

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents. Comments or suggestions should be addressed to:

Commander
USA Concepts Analysis Agency
ATTN: Director of Methodology,
Resources and Computation
8120 Woodmont Avenue
Bethesda, MD 20014

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: CAA-TP-76-3	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER: Final rept.
4. TITLE (and Subtitle): An Application of the Multivariate Analysis System to TAGS		5. TYPE OF REPORT & PERIOD COVERED: Technical Paper (Final)
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): ~~Dr.~~ Daniel A. Nussbaum, Resource Analysis Group, MRCD		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: US Army Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20014		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: ~~6992F~~
11. CONTROLLING OFFICE NAME AND ADDRESS: US Army Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20014		12. REPORT DATE: December 1976
		13. NUMBER OF PAGES: 82 (82p.)
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): US Army Concepts Analysis Agency, 8120 Woodmont Avenue, Bethesda, MD 20014		15. SECURITY CLASS. (of this report): UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Air-Ground Interactions
Multivariate Techniques
Principal Components
Multiple Regression
Canonical Correlation
Fractional factoral design

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Theater Air Ground Study (TAGS) is a highly aggregated computer model of air-ground warfare. This report demonstrates the applicability of the Multivariate Analysis System to the mission of analyzing data obtained from the model. Analyses are performed which encompass the topics of multiple regression, principal components, and canonical correlation. Results are displayed in the form of equations, tables and graphs.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 390 996

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

bpg

AN APPLICATION OF THE MULTIVARIATE
ANALYSIS SYSTEM TO TAGS

DECEMBER 1976

PREPARED BY

METHODOLOGY, RESOURCES AND COMPUTATION DIRECTORATE

US ARMY CONCEPTS ANALYSIS AGENCY
8120 WOODMONT AVENUE
BETHESDA, MARYLAND 20014

AN APPLICATION OF THE MULTIVARIATE ANALYSIS SYSTEM TO TAGS

SUMMARY

1. Purpose. The purpose of this paper is to demonstrate and document the applicability of the Multivariate Analysis System to the mission of analyzing data obtained from a theater level model impacting on air-ground interactions.

2. Objective. The objective is to analyze multivariate data generated by the exercise of a theater level model. The analysis includes means, standard deviations, multiple regression, principal components, rotation of axes, and canonical correlation.

3. Discussion. A fractional factorial experiment with each factor having three levels was chosen for the analysis. The base value for each input variable was taken from a benchmark case provided by a RAND listing of the TAGS Model. Model design, factor inputs, outputs, calculations, problems, and observations are discussed in the paper.

4. Observations. Based upon the analytic work developed in this paper, the following observations are provided. The problem of analyzing complex multivariate data is amenable to the techniques which form the Multivariate Analysis System. Analysis of correlation coefficients provides a means for reducing redundancy in the data. Multiple regression yields mathematical functions which possess good predictive capabilities over the range of the original data. Principal component analysis, together with the rotation of axes provides a technique for viewing data in a different perspective with corresponding insights into the data. The results were consistent with expectations.

TABLE OF CONTENTS

FIGURES

Figure | Page

TABLES

Table

AN APPLICATION OF THE MULTIVARIATE ANALYSIS SYSTEM TO TAGS

1. Purpose. The purpose of this report is to demonstrate applicability of the Multivariate Analysis System to the mission of analyzing data obtained from a theater level model impacting on air-ground interactions. To this end, analyses will encompass the following topics:

Basic statistics	Principal components
Multiple regression	Canonical Correlation

2. Background. Theater Air-Ground Study (TAGS) is a highly aggregated computer model of air-ground warfare. Developed by D.E. Emerson of RAND in the early 1950's, the model has since undergone extensive revisions, modifications, and enhancements. The present design allows for consideration of reserve forces, individual and unit replacements, repair capabilities, terrain, Surface to Air Missile (SAM) and counter-SAM, and prepared defenses in addition to earlier options.

3. Design of Experiment

a. Assessing the Effect. Initially six input variables and 13 output variables were chosen for study because of their impact on air-ground interactions. Table I contains a list of the initially chosen variables. In order to assess the effect of the six impact variables on the 13 output variables, a sequence of runs of the TAGS Model was undertaken.

b. Levels. The six input variables range through three levels: a base value; a high value, given by 110 percent of the base value; and a low value, given by 90 percent of the base value. The base value for each input variable was taken from a benchmark case provided by a RAND listing of the TAGS Model. Table II displays the various values for the six input variables.

c. Factorial Design. Each level of the six input variables appears in combination with each level of all other input variables. Consequently, a full design contains 3^6 or 729 design points. Because of the large number of TAGS runs required for a full design a more efficient design was sought. Investigation revealed that a 1/9 th replicate requiring 81 runs could be performed without extensive loss of information. Therefore, the experiment was

TABLE I, Initially Chosen Variables (continued on next page)

Name		Description
Inputs		
1.	BDIV	Number of Blue division slices.
2.	BAC	Number of Blue aircraft.
3.	BSHLT	Number of Blue aircraft shelters.
4.	RDIV	Number of Red division slices.
5.	RAC	Number of Red aircraft.
6.	RSHLT	Number of Red aircraft shelters.
Outputs		
1.	FEBA	FEBA (positive for Blue advance; negative for Red advance)
2.	BFTRLOSS	Number of Blue fighter-type aircraft lost.
3.	BATTLOSS	Number of Blue attacker-type aircraft lost.
4.	BBMRLOSS	Number of Blue bomber-type aircraft lost.
5.	BDIVLOSSTOGR	Number of Blue divisions lost to enemy ground forces.
6.	BDIVLOSS2AI	Number of Blue divisions lost to enemy air forces.
7.	BTOTGDLOSS	Number of Blue divisions lost (total).
8.	RFTRLOSS	Number of Red fighter-type aircraft lost.
9.	RATTLOSS	Number of Red attacker-type aircraft lost.
10.	RBMRLOSS	Number of Red bomber-type aircraft lost.

TABLE I, Initially Chosen Variables (concluded)

Name		Description
Outputs		
11.	RDIVLOSS2GR	Number of Red divisions lost to enemy ground forces.
12.	RDIVLOSS2AI	Number of Red divisions lost to enemy air forces.
13.	RTOTGDLOSS	Number of Red divisions lost (total).

TABLE II, Input Values

Input Variables	Low	Base	High
1. Blue divisions	15.9	19.9	21.9
2. Blue aircraft	1997.0	2496.0	2746.0
3. Blue aircraft shelters	818.0	1023.0	1126.0
4. Red divisions	17.2	21.5	23.6
5. Red aircraft	2440.0	3050.0	3355.0
6. Red aircraft shelters	1171.0	1464.0	1610.0

fractionated, and a 1/9 x 3^6 fractional factorial design was used. Table III identifies the 81 design points. The low, base, and high values are represented by 0, 1, and 2, respectively. The position of the six digits identifies the six input variables. For example, the first design point (000000) is the low value of all six input variables. Design point 28 (100110) is the low value of input variables 2, 3, and 6 and the base value of input variables 1, 4, and 5. The other design points are read similarly. The design may be found in "Fractional Factorial Experiments Designs for Factor at Three Levels," US Department of Commerce, National Bureau of Standards, Applied Mathematics Series No. 54.

4. Reduction of Variables. No loss of information ensues if the output variables BTOTGDLOSS and RTOTGDLOSS are deleted from the analysis. This is true because the information contained in BTOTGDLOSS can be reconstructed by the variables BDIVLOSS2AI and BDIVLOSSTOGR. Indeed BTOTGDLOSS = BDIVLOSS2AI + BDIVLOSSTOGD. A similar argument holds for RTOTGDLOSS. The data base generated this way appears in Table IV.

5. Graphical Representation of the Data. Viewing graphical portrayals of the input and output variables is instructive. These graphs were generated by the Multi-Optioned Interactive Display and Analytic System (MIDAS), which is not part of MAS. Figures I-VI represent the six input variables. The three levels for these variables can be discerned clearly by distinctive clusters in the graphs. Similarly, Figures VII - XVII allow the analyst to observe the range of the 11 output variables and do preliminary clustering. For example, Figure VII indicates a clustering of cases into two clusters, wherein the FEBA gives a measurement of (1) approximately -100 kilometers, and (2) a range of -50 to -70 kilometers. Figure XV shows clearly that losses of red bombers fall into three distinct clusters: those representing losses of approximately 480 aircraft, 600 aircraft and approximately 650 aircraft, respectively.

6. Computations

a. Basic Statistics

(1) Means and Standard Deviations. Computations were made of the means and standard deviations of variables in the data base, as listed in Table IV. Table V lists the means and standard deviations for the six input variables and 11 output variables.

(2) Correlations. The Multivariate Analysis System computes and stores all pairwise product moment correlation coefficients. For 17 variables, there are 289 such pairwise correlation coefficients. However, since the correlation coefficient

TABLE III, Design of Experiment
(1/9 Replicate of a 3^6 Factorial Experiment)

1.	000000	28.	100110	55.	200220
2.	112122	29.	212202	56.	012012
3.	221211	30.	021021	57.	121101
4.	000121	31.	100201	58.	200011
5.	112210	32.	212020	59.	012100
6.	221002	33.	021112	60.	121222
7.	000212	34.	100022	61.	200102
8.	112001	35.	212111	62.	012221
9.	221120	36.	021200	63.	121010
10.	201210	37.	001020	64.	101100
11.	010002	38.	110112	65.	210222
12.	122121	39.	222201	66.	022011
13.	201001	40.	001111	67.	101221
14.	010120	41.	110200	68.	210010
15.	122212	42.	222022	69.	022102
16.	201122	43.	001202	70.	101012
17.	010211	44.	110021	71.	210101
18.	122000	45.	222110	72.	022220
19.	102120	46.	202200	73.	002010
20.	211212	47.	011022	74.	111102
21.	020001	48.	120111	75.	220221
22.	102211	49.	202021	76.	002101
23.	211000	50.	011110	77.	111220
24.	020122	51.	120202	78.	220012
25.	102002	52.	202112	79.	002222
26.	211121	53.	011201	80.	111011
27.	020210	54.	120020	81.	220100

Table IV, Data Base (continued on next page)

RUN NO.	BLUE DIVS FEBA	BLUE A/C B FTR LOSS R FTR LOSS	B A/C SHELT B ATK LOSS R ATK LOSS	RED DIVS B BMR LOSS R BMR LOSS	RED A/C B GD/GD LSS R GD/GD LSS	RED A/C SHLT B AIR/GD LSS R AIR/GD LS
	15.93	1997.00	818.00	17.19	2440.00	1171.00
1	-61.20	537.00	708.00	455.00	7.83	15.30
		1386.00	339.00	480.00	11.38	4.94
	19.91	2496.00	1126.00	21.49	3355.00	1610.00
2	-61.70	813.00	923.00	606.00	7.53	15.20
		1752.00	440.00	658.00	11.26	3.99
	21.90	2746.00	1023.00	23.64	3050.00	1464.00
3	-52.60	859.00	961.00	596.00	7.44	15.50
		1836.00	438.00	602.00	9.74	6.85
	15.93	1997.00	818.00	21.49	3355.00	1464.00
4	-123.10	745.00	784.00	502.00	9.44	15.00
		1284.00	332.00	650.00	13.68	1.34
	19.91	2496.00	1126.00	23.64	3050.00	1171.00
5	-54.50	776.00	799.00	544.00	7.79	17.30
		1851.00	440.00	603.00	10.76	6.63
	21.90	2746.00	1023.00	17.19	2440.00	1610.00
6	-49.20	774.00	737.00	536.00	7.36	18.50
		1579.00	373.00	483.00	8.93	9.60
	15.93	1997.00	818.00	23.64	3050.00	1610.00
7	-100.80	709.00	784.00	502.00	9.35	15.50
		1254.00	323.00	594.00	14.76	1.75
	19.91	2496.00	1126.00	17.19	2440.00	1464.00
8	-53.00	728.00	712.00	502.00	7.73	19.90
		1549.00	363.00	483.00	10.50	8.46
	21.90	2746.00	1023.00	21.49	3355.00	1171.00
9	-58.50	876.00	880.00	607.00	7.49	16.80
		2033.00	484.00	662.00	10.33	5.55
	21.90	1997.00	1023.00	23.64	3050.00	1171.00
10	-90.00	659.00	784.00	496.00	8.57	15.70
		1517.00	386.00	598.00	12.89	2.82
	15.93	2496.00	818.00	17.19	2440.00	1610.00
11	-56.10	772.00	756.00	525.00	7.71	17.00
		1493.00	353.00	482.00	9.77	7.24
	19.91	2746.00	1126.00	21.49	3355.00	1464.00
12	-58.80	872.00	914.00	621.00	7.65	15.40
		1949.00	472.00	661.00	10.41	5.05
	21.90	1997.00	1023.00	17.19	2440.00	1464.00
13	-70.70	620.00	715.00	455.00	8.41	16.90
		1354.00	334.00	480.00	11.90	5.06
	15.93	2496.00	818.00	21.49	3355.00	1171.00
14	-68.20	845.00	917.00	604.00	8.28	16.10
		1818.00	452.00	659.00	12.30	3.81
	19.91	2746.00	1126.00	23.64	3050.00	1610.00
15	-57.30	852.00	868.00	599.00	7.90	17.60
		1811.00	434.00	602.00	10.71	6.38
	21.90	1997.00	1023.00	21.49	3355.00	1610.00
16	-98.40	699.00	784.00	502.00	9.18	16.10
		1347.00	342.00	652.00	14.55	1.57
	15.93	2496.00	818.00	23.64	3050.00	1464.00
17	-61.20	825.00	890.00	589.00	7.95	16.20
		1673.00	415.00	600.00	11.66	4.55
	19.91	2746.00	1126.00	17.19	2440.00	1171.00
18	-44.40	695.00	682.00	509.00	7.33	19.50
		1579.00	375.00	483.00	9.01	10.56

Table IV, Data Base (continued on next page)

RUN NO.	BLUE DIVS FEBA	BLUE A/C	B A/C SHELT	RED DIVS	RED A/C	RED A/C SHLT
		B FTR LOSS	B ATK LOSS	B BMR LOSS	B GD/GD LSS	B AIR/GD LSS
		R FTR LOSS	R ATK LOSS	R BMR LOSS	R GD/GD LSS	R AIR/GD LS
	19.91	1997.00	1125.00	21.49	3355.00	1171.00
19	-99.10	567.00	784.00	502.00	9.36	16.70
		1545.00	393.00	655.00	14.78	2.00
	21.90	2496.00	1023.00	23.64	3050.00	1610.00
20	-62.20	892.00	878.00	581.00	7.69	16.00
		1580.00	416.00	600.00	11.10	5.00
	15.93	2746.00	818.00	17.19	2440.00	1464.00
21	-51.80	787.00	743.00	541.00	7.45	19.00
		1579.00	373.00	483.00	9.38	9.62
	19.91	1997.00	1126.00	23.64	3050.00	1464.00
22	-97.20	650.00	784.00	499.00	9.33	16.80
		1445.00	374.00	597.00	14.17	2.65
	21.90	2496.00	1023.00	17.19	2440.00	1171.00
23	-50.70	702.00	683.00	490.00	7.38	19.50
		1579.00	374.00	483.00	9.75	8.95
	15.93	2746.00	818.00	21.49	3355.00	1610.00
24	-63.80	918.00	979.00	657.00	8.03	16.10
		1809.00	452.00	659.00	11.70	4.44
	19.91	1997.00	1126.00	17.19	2440.00	1610.00
25	-65.30	611.00	713.00	453.00	8.33	16.90
		1356.00	333.00	480.00	11.80	5.18
	21.90	2495.00	1023.00	21.49	3355.00	1464.00
26	-80.00	824.00	921.00	605.00	8.25	15.40
		1773.00	444.00	658.00	11.50	3.94
	15.93	2746.00	818.00	23.64	3050.00	1171.00
27	-54.00	370.00	837.00	588.00	7.67	17.50
		1900.00	448.00	603.00	10.21	7.29
	19.91	1997.00	818.00	21.49	3050.00	1171.00
28	-98.00	594.00	784.00	502.00	9.44	16.80
		1412.00	365.00	596.00	14.65	2.16
	21.90	2496.00	1126.00,	23.64	2440.00	1610.00
29	-55.90	737.00	724.00	507.00	7.28	17.70
		1531.00	359.00	483.00	9.60	8.11
	15.93	2745.00	1023.00	17.19	3355.00	1464.00
30	-62.70	885.00	928.00	629.00	7.93	16.80
		1925.00	468.00	660.00	11.16	5.65
	19.31	1997.00	818.00	23.64	2440.00	1464.00
31	-65.40	646.00	742.00	470.00	8.45	16.60
		1306.00	327.00	479.00	12.36	4.29
	21.90	2496.00	1126.00	17.19	3355.00	1171.00
32	-61.20	800.00	868.00	576.00	7.74	15.00
		1935.00	470.00	661.00	10.97	5.09
	15.93	2746.00	1023.00	21.49	3050.00	1610.00
33	-56.00	863.00	881.00	606.00	7.68	17.10
		1792.00	432.00	601.00	10.57	5.58
	19.91	1997.00	818.00	17.19	3355.00	1610.00
34	-114.30	751.00	784.00	502.00	9.23	15.10
		1743.00	324.00	649.00	13.86	1.30
	21.90	2496.00	1126.00	21.49	3050.00	1464.00
35	-62.00	784.00	845.00	564.00	8.05	17.30
		1751.00	426.00	601.00	11.63	5.68
	15.93	2746.00	1023.00	23.64	2440.00	1171.00
36	-53.30	709.00	699.00	514.00	7.49	19.30
		1579.00	375.00	483.00	8.86	10.51

Table IV, Data Base (continued on next page)

RUN NO.	BLUE DIVS FEBA	BLUE A/C B FTR LOSS R FTR LOSS	B A/C SHELT B ATK LOSS R ATK LOSS	RED DIVS B BMR LOSS R BMR LOSS	RED A/C B GD/GD LSS R GD/GD LSS	RED A/C SHLT B AIR/GD LSS R AIR/GD LS
	19.91	1997.00	1125.00	21.49	3355.00	1171.00
19	-99.10	667.00	784.00	502.00	9.36	16.70
		1545.00	393.00	655.00	14.78	2.00
	21.90	2496.00	1023.00	23.64	3050.00	1610.00
20	-62.20	892.00	878.00	581.00	7.69	16.00
		1580.00	416.00	600.00	11.10	5.00
	15.93	2746.00	818.00	17.19	2440.00	1464.00
21	-51.80	787.00	743.00	541.00	7.45	19.00
		1579.00	373.00	483.00	9.38	9.62
	19.91	1997.00	1126.00	23.64	3050.00	1464.00
22	-97.20	650.00	784.00	499.00	9.33	16.80
		1445.00	374.00	597.00	14.17	2.65
	21.90	2495.00	1023.00	17.19	2440.00	1171.00
23	-50.70	702.00	683.00	490.00	7.38	19.50
		1579.00	374.00	483.00	9.75	8.85
	15.93	2746.00	818.00	21.49	3355.00	1610.00
24	-63.80	918.00	979.00	657.00	8.03	16.10
		1809.00	452.00	659.00	11.70	4.44
	19.91	1997.00	1126.00	17.19	2440.00	1610.00
25	-65.30	611.00	713.00	453.00	8.33	16.90
		1356.00	333.00	480.00	11.80	5.18
	21.90	2495.00	1023.00	21.49	3355.00	1464.00
26	-80.00	824.00	921.00	605.00	8.25	15.40
		1773.00	444.00	658.00	11.50	3.94
	15.93	2746.00	818.00	23.64	3050.00	1171.00
27	-54.00	970.00	837.00	588.00	7.67	17.50
		1900.00	448.00	603.00	10.21	7.29
	19.91	1997.00	818.00	21.49	3050.00	1171.00
28	-98.00	594.00	784.00	502.00	9.44	16.80
		1412.00	365.00	596.00	14.65	2.16
	21.90	2496.00	1126.00,	23.64	2440.00	1610.00
29	-55.90	737.00	724.00	507.00	7.28	17.70
		1531.00	359.00	483.00	9.60	8.11
	15.93	2746.00	1023.00	17.19	3355.00	1464.00
30	-62.70	885.00	928.00	629.00	7.93	16.80
		1925.00	468.00	660.00	11.16	5.65
	19.91	1997.00	818.00	23.64	2440.00	1464.00
31	-65.40	646.00	742.00	470.00	8.45	15.60
		1306.00	327.00	479.00	12.36	4.29
	21.90	2496.00	1126.00	17.19	3355.00	1171.00
32	-61.70	800.00	868.00	576.00	7.74	15.00
		1935.00	470.00	661.00	10.97	5.09
	15.93	2746.00	1023.00	21.49	3050.00	1610.00
33	-56.00	863.00	881.00	606.00	7.68	17.10
		1792.00	432.00	601.00	10.57	6.58
	19.91	1997.00	819.00	17.19	3355.00	1610.00
34	-114.30	751.00	784.00	502.00	9.23	15.10
		1743.00	324.00	649.00	13.86	1.30
	21.90	2496.00	1126.00	21.49	3050.00	1464.00
35	-67.00	784.00	845.00	564.00	8.05	17.30
		1751.00	426.00	601.00	11.57	5.68
	15.93	2746.00	1023.00	23.64	2440.00	1171.00
36	-53.30	709.00	689.00	514.00	7.49	19.30
		1579.00	375.00	483.00	8.86	10.51

Table IV, Data Base (continued on next page)

	BLUE DIVS	BLUE A/C	B A/C SHELT	RED DIVS	RED A/C	RED A/C SHLT
RUN	FEBA	B FTR LOSS	B ATK LOSS	B BMR LOSS	B GD/GD LSS	B AIR/GD LSS
NO.		R FTR LOSS	R ATK LOSS	R BMR LOSS	R GD/GD LSS	R AIR/GD LS
	15.93	1997.00	1023.00	17.19	3355.00	1171.00
37	-99.00	686.00	784.00	502.00	9.23	15.00
		1491.00	377.00	654.00	14.25	1.79
	19.91	2496.00	818.00	21.49	3050.00	1610.00
38	-61.20	830.00	907.00	598.00	7.79	15.80
		1618.00	407.00	599.00	11.56	4.27
	21.90	2746.00	1126.00	23.64	2440.00	1464.00
39	-48.70	742.00	716.00	526.00	7.31	19.10
		1573.00	375.00	483.00	9.32	9.81
	15.93	1997.00	1023.00	21.49	3050.00	1464.00
40	-90.00	666.00	784.00	501.00	8.68	15.80
		1401.00	364.00	596.00	13.46	2.38
	19.91	2496.00	818.00	23.64	2440.00	1171.00
41	-51.00	735.00	704.00	501.00	7.46	18.00
		1573.00	370.00	483.00	9.81	8.26
	21.90	2746.00	1126.00	17.19	3355.00	1610.00
42	-66.80	877.00	935.00	632.00	7.87	16.40
		1899.00	465.00	660.00	10.80	5.67
	15.93	1997.00	1023.00	23.64	2440.00	1610.00
43	-61.10	523.00	725.00	460.00	7.84	15.90
		1334.00	330.00	479.00	11.17	4.81
	19.91	2496.00	818.00	17.19	3355.00	1464.00
44	-85.90	857.00	929.00	617.00	8.34	15.90
		1672.00	427.00	657.00	12.80	3.19
	21.90	2746.00	1126.00	21.49	3050.00	1171.00
45	-52.60	826.00	799.00	567.00	7.30	16.70
		1945.00	456.00	604.00	8.71	8.07
	21.90	1997.00	1126.00	23.64	2440.00	1171.00
46	-67.00	601.00	668.00	433.00	8.32	17.80
		1449.00	347.00	481.00	11.91	5.92
	15.93	2496.00	1023.00	17.19	3355.00	1610.00
47	-67.20	828.00	928.00	612.00	7.96	15.50
		1706.00	432.00	658.00	11.90	3.65
	19.91	2746.00	818.00	21.49	3050.00	1464.00
48	-59.00	882.00	888.00	611.00	7.58	16.80
		1797.00	433.00	602.00	10.54	6.30
	21.90	1997.00	1126.00	17.19	3355.00	1464.00
49	-114.10	675.00	784.00	502.00	9.40	15.80
		1441.00	365.00	654.00	14.08	1.79
	15.93	2496.00	1023.00	21.49	3050.00	1171.00
50	-60.10	788.00	814.00	551.00	8.01	17.50
		1830.00	437.00	602.00	11.38	6.14
	19.91	2746.00	818.00	23.64	2440.00	1610.00
51	-52.80	804.00	755.00	547.00	7.46	18.90
		1571.00	368.00	483.00	9.46	9.41
	21.90	1997.00	1126.00	21.49	3050.00	1610.00
52	-97.20	653.00	784.00	501.00	9.14	16.10
		1393.00	363.00	596.00	13.66	2.46
	15.93	2496.00	1023.00	23.64	2440.00	1464.00
53	-55.80	741.00	722.00	507.00	7.28	17.60
		1578.00	361.00	483.00	9.61	8.07
	19.91	2746.00	818.00	17.19	3355.00	1171.00
54	-62.00	902.00	912.00	624.00	7.98	15.80
		1984.00	477.00	561.00	11.13	5.69

Table IV, Data Base (continued on next page)

RUN NO.	BLUE DIVS FEBA	BLUE A/C B FTR LOSS / R FTR LOSS	B A/C SHELT B ATK LOSS / R ATK LOSS	RED DIVS B BMR LOSS / R BMR LOSS	RED A/C R GD/GD LSS / R GD/GD LSS	RED A/C SHLT B AIR/GD LSS / R AIR/GD LS
	21.90	1997.00	818.00	23.64	3355.00	1171.00
55	-97.80	732.00	784.00	507.00	9.20	16.00
		1371.00	349.00	652.00	14.65	1.44
	15.93	2495.00	1126.00	17.19	3050.00	1610.00
56	-61.00	788.00	864.00	573.00	7.69	16.30
		1705.00	420.00	600.00	10.95	5.39
	19.91	2746.00	1023.00	21.49	2440.00	1464.00
57	-46.60	757.00	725.00	531.00	7.73	18.40
		1579.00	375.00	483.00	8.73	9.71
	21.90	1997.00	818.00	17.19	3050.00	1464.00
58	-94.90	704.00	784.00	502.00	8.95	16.20
		1303.00	336.00	594.00	14.35	1.85
	15.93	2496.00	1126.00	21.49	2440.00	1171.00
59	-50.60	686.00	674.00	485.00	7.40	18.70
		1579.00	375.00	483.00	9.70	9.05
	19.91	2746.00	1023.00	23.64	3355.00	1610.00
60	-58.50	890.00	949.00	640.00	7.31	15.50
		1872.00	461.00	660.00	10.38	5.21
	21.90	1997.00	818.00	21.49	2440.00	1610.00
61	-64.80	650.00	752.00	474.00	8.14	15.70
		1284.00	323.00	478.00	12.14	4.13
	15.93	2495.00	1126.00	23.64	3355.00	1464.00
62	-73.80	809.00	908.00	597.00	8.25	16.10
		1818.00	452.00	659.00	11.87	4.33
	19.91	2746.00	1023.00	17.19	3050.00	1171.00
63	-54.90	842.00	811.00	574.00	7.56	17.20
		1931.00	453.00	604.00	9.44	7.78
	19.91	1997.00	1023.00	21.49	2440.00	1171.00
64	-68.00	613.00	680.00	440.00	8.37	17.50
		1429.00	344.00	481.00	11.93	5.64
	21.90	2496.00	818.00	23.64	3355.00	1610.00
65	-95.70	862.00	929.00	620.00	9.07	16.60
		1508.00	417.00	656.00	13.69	2.94
	15.93	2746.00	1126.00	17.19	3050.00	1464.00
66	-55.40	847.00	848.00	589.00	7.68	17.30
		1854.00	440.00	603.00	10.08	7.31
	19.91	1997.00	1023.00	23.64	3355.00	1464.00
67	-114.40	695.00	784.00	507.00	9.41	15.60
		1391.00	352.00	653.00	14.06	1.63
	21.90	2496.00	818.00	17.19	3050.00	1171.00
68	-58.00	914.00	845.00	568.00	7.53	16.20
		1787.00	431.00	601.00	10.93	5.35
	15.93	2746.00	1126.00	21.49	2440.00	1610.00
69	-46.40	759.00	728.00	531.00	7.30	18.40
		1579.00	374.00	483.00	8.77	9.72
	19.91	1997.00	1023.00	17.19	3050.00	1610.00
70	-90.00	670.00	784.00	502.00	8.79	16.10
		1351.00	351.00	595.00	13.98	7.21
	21.90	2496.00	818.00	21.49	2440.00	1464.00
71	-55.00	765.00	744.00	519.00	7.36	17.30
		1514.00	357.00	482.00	10.07	7.34
	15.93	2746.00	1126.00	23.64	3355.00	1171.00
72	-56.50	863.00	865.00	599.00	7.58	17.70
		2055.00	487.00	663.00	10.39	5.89

Table IV, Data Base (concluded)

RUN NO.	BLUE DIVS FEBA	BLUE A/C B FTR LOSS R FTR LOSS	B A/C SHELT B ATK LOSS R ATK LOSS	RED DIVS B BMR LOSS R BMR LOSS	RED A/C B GD/GD LSS R GD/GD LSS	RED A/C SHLT B AIR/GD LSS R AIR/GD LS
	15.93	1997.00	1126.00	17.19	3050.00	1171.00
73	-90.00	643.00	784.00	497.00	8.99	16.80
		1565.00	394.00	598.00	13.67	3.15
	19.91	2496.00	1023.00	21.49	2440.00	1510.00
74	-54.80	749.00	734.00	513.00	7.25	17.50
		1519.00	357.00	483.00	9.62	7.88
	21.90	2746.00	818.00	23.64	3355.00	1464.00
75	-63.00	913.00	960.00	647.00	8.39	17.40
		1869.00	461.00	660.00	12.66	4.79
	15.93	1997.00	1126.00	21.49	2440.00	1464.00
76	-67.80	608.00	703.00	448.00	8.31	17.10
		1375.00	336.00	480.00	11.84	5.35
	19.91	2496.00	1023.00	23.64	3355.00	1171.00
77	-63.70	815.00	884.00	585.00	7.95	16.70
		1903.00	465.00	660.00	11.67	4.61
	21.90	2746.00	818.00	17.19	3050.00	1610.00
78	-60.30	887.00	909.00	621.00	7.52	16.50
		1750.00	427.00	601.00	10.63	5.90
	15.93	1997.00	1126.00	23.64	3355.00	1610.00
79	-114.40	679.00	784.00	507.00	9.41	15.70
		1394.00	353.00	653.00	14.06	1.71
	19.91	2496.00	1023.00	17.19	3050.00	1464.00
80	-60.50	797.00	860.00	572.00	7.97	16.90
		1727.00	423.00	600.00	11.57	5.35
	21.90	2746.00	818.00	21.49	2440.00	1171.00
81	-52.80	736.00	704.00	523.00	7.47	19.10
		1579.00	375.00	483.00	8.91	10.22

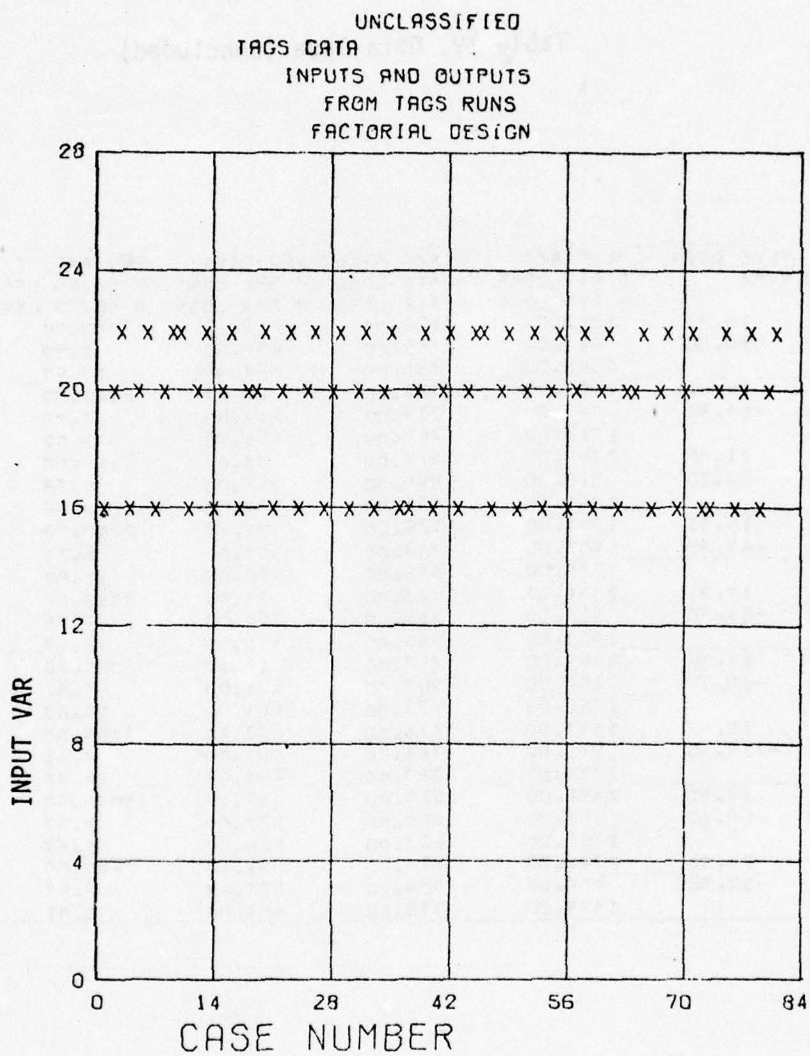


Figure I, Blue Division

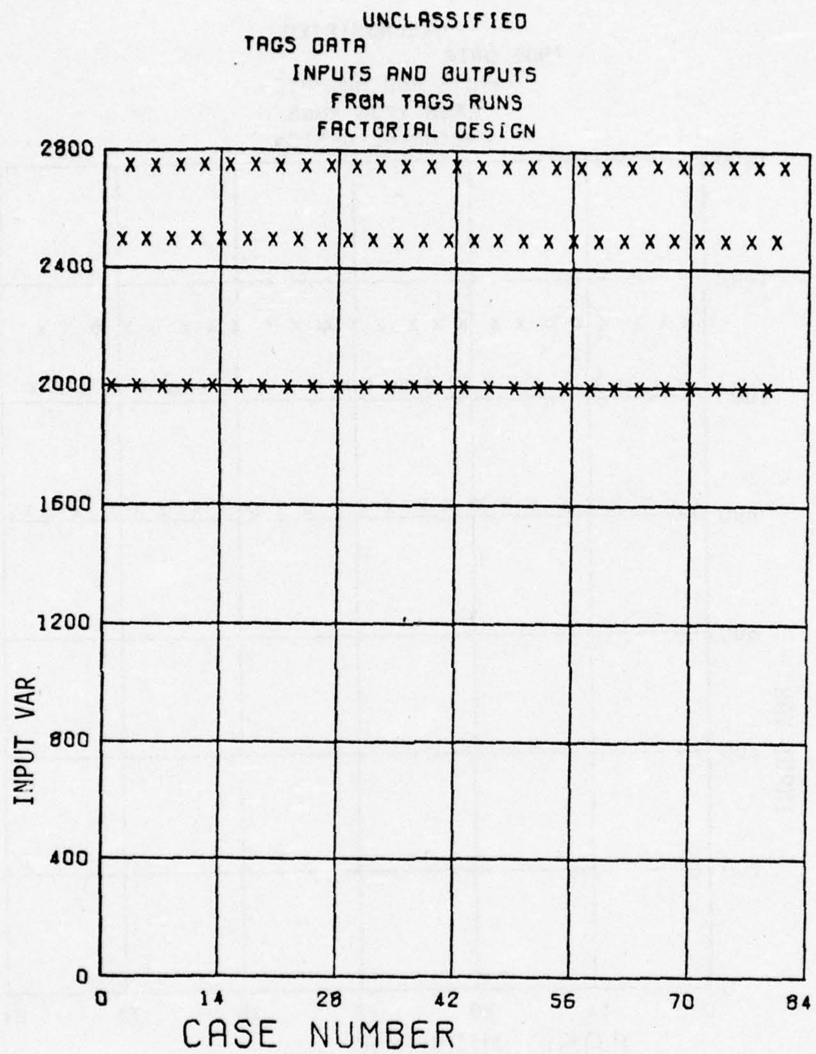


Figure II, Blue Total Aircraft

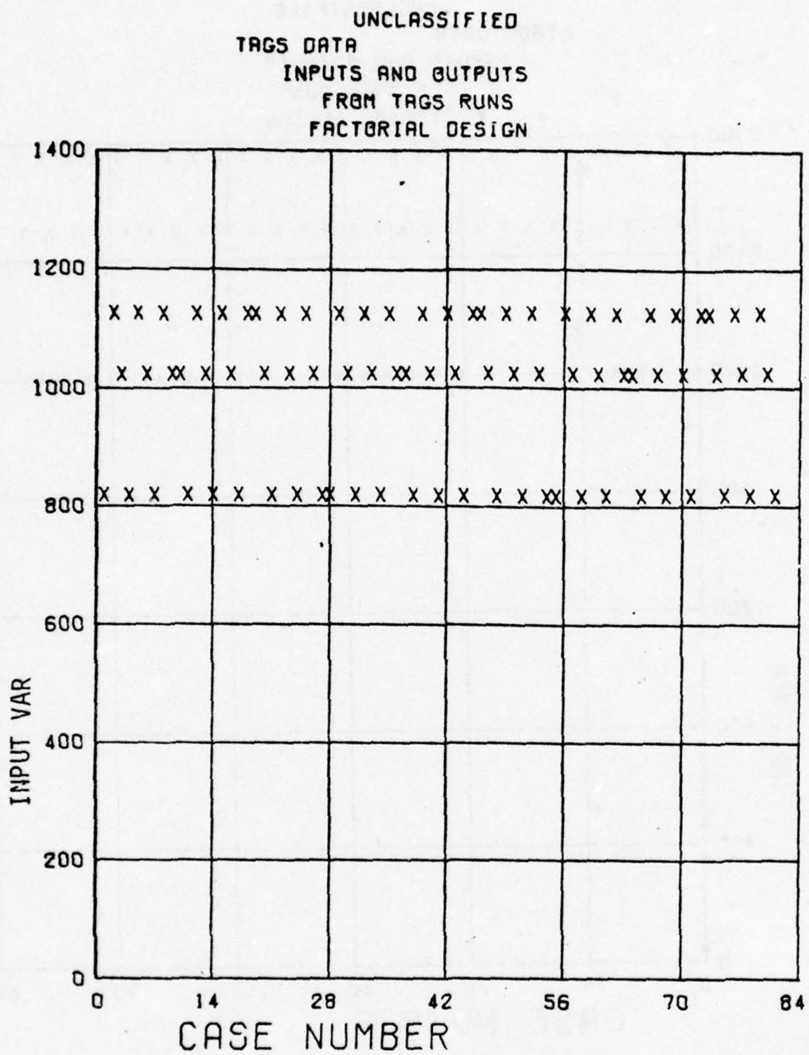


Figure III, Blue Aircraft Shelters

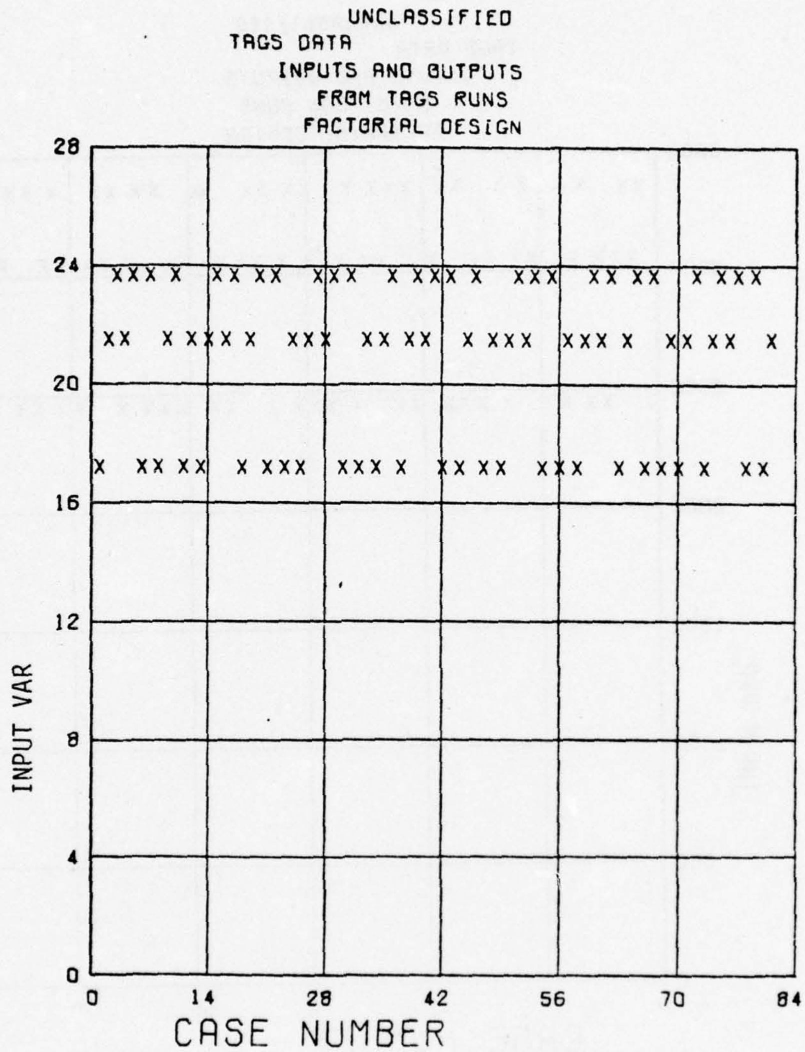


Figure IV, Red Divisions

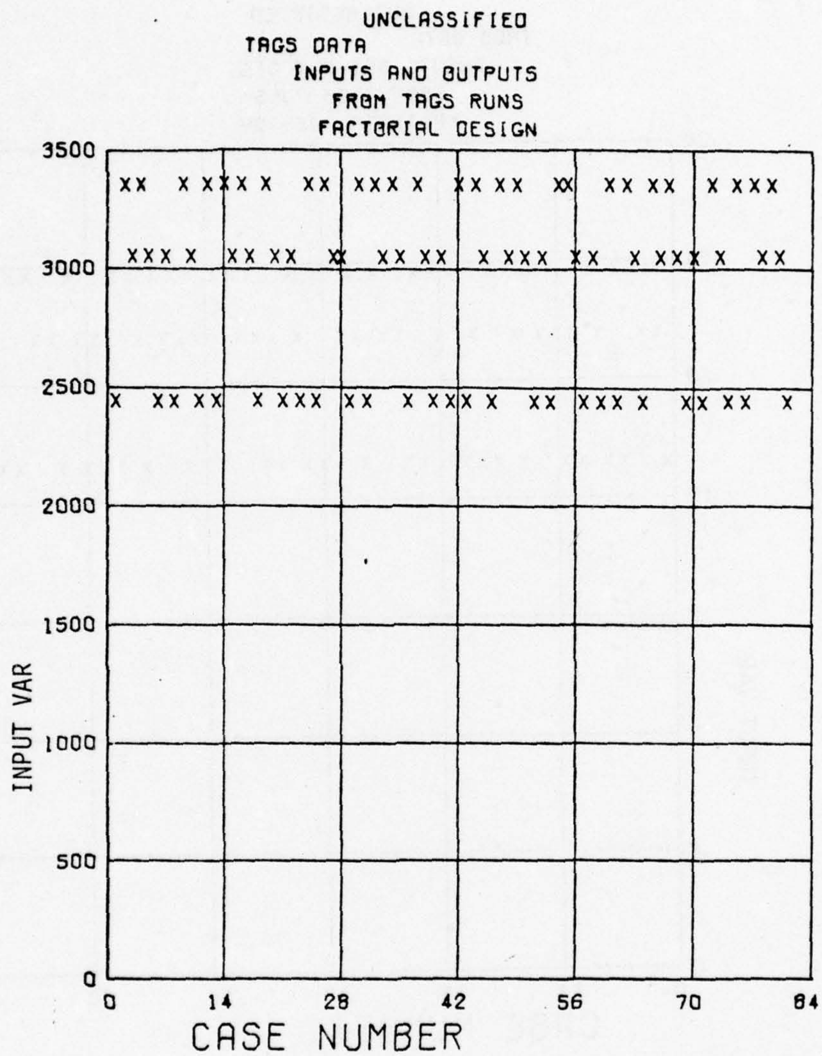


Figure V, Red Total Aircraft

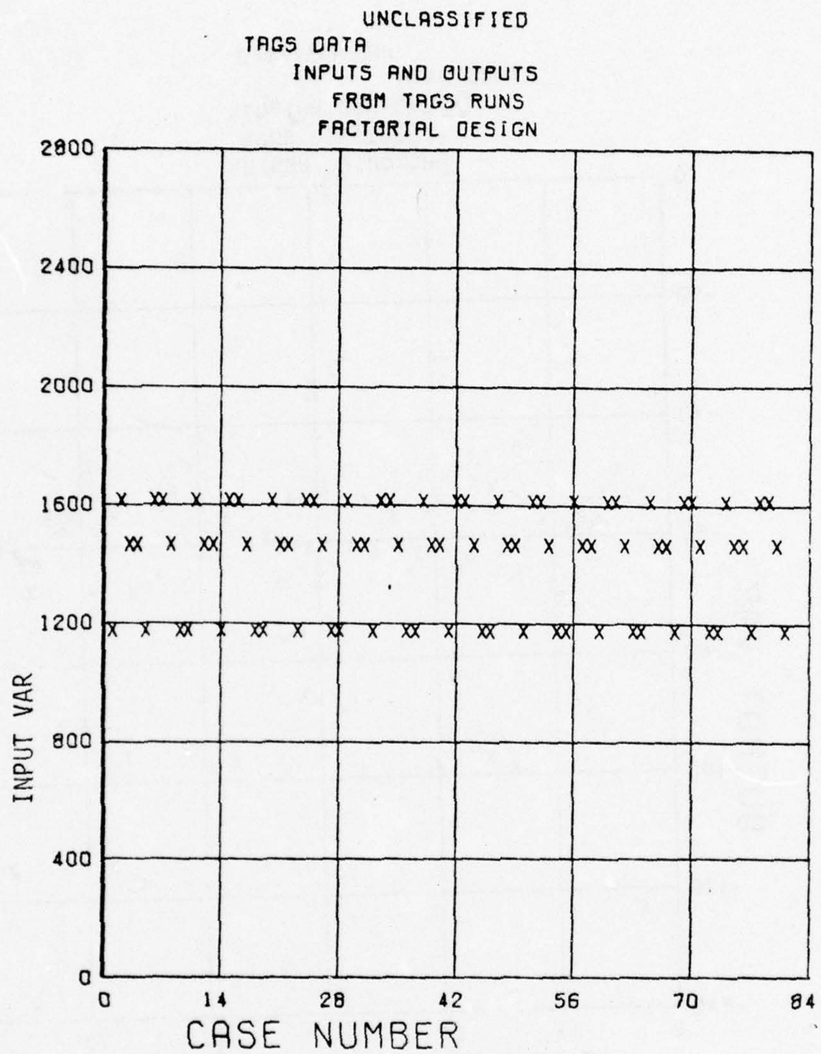


Figure VI, Red Aircraft Shelters

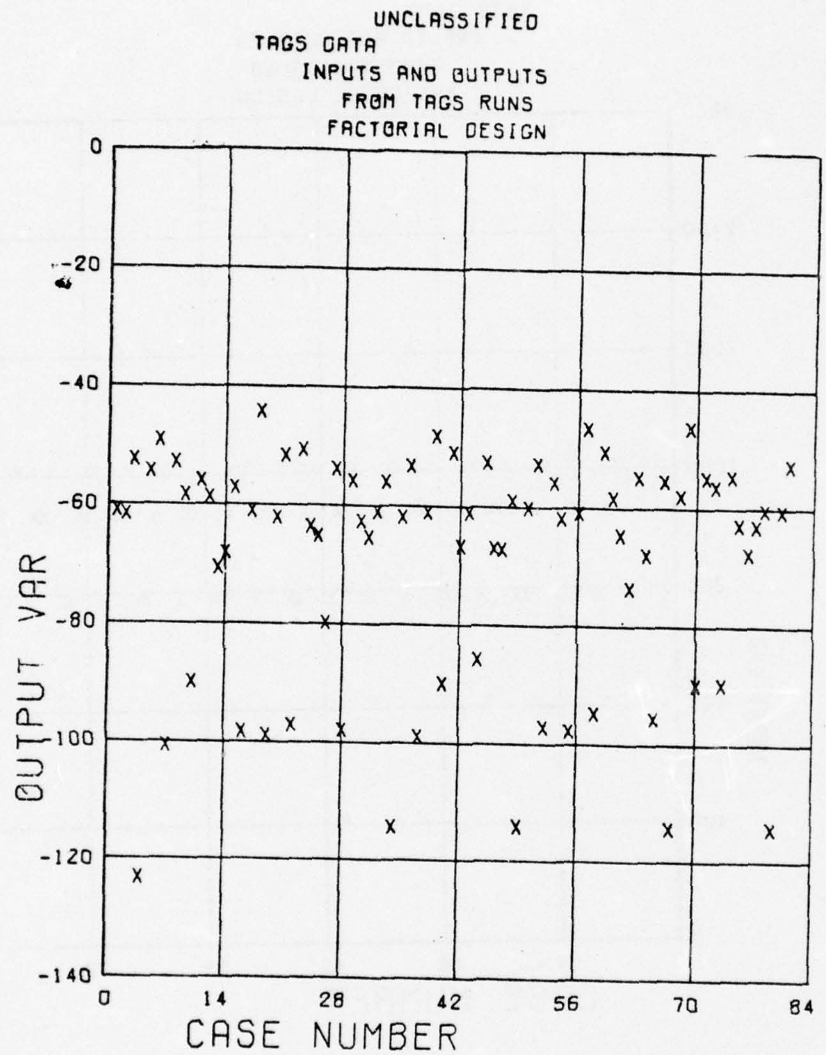


Figure VII, FEBA

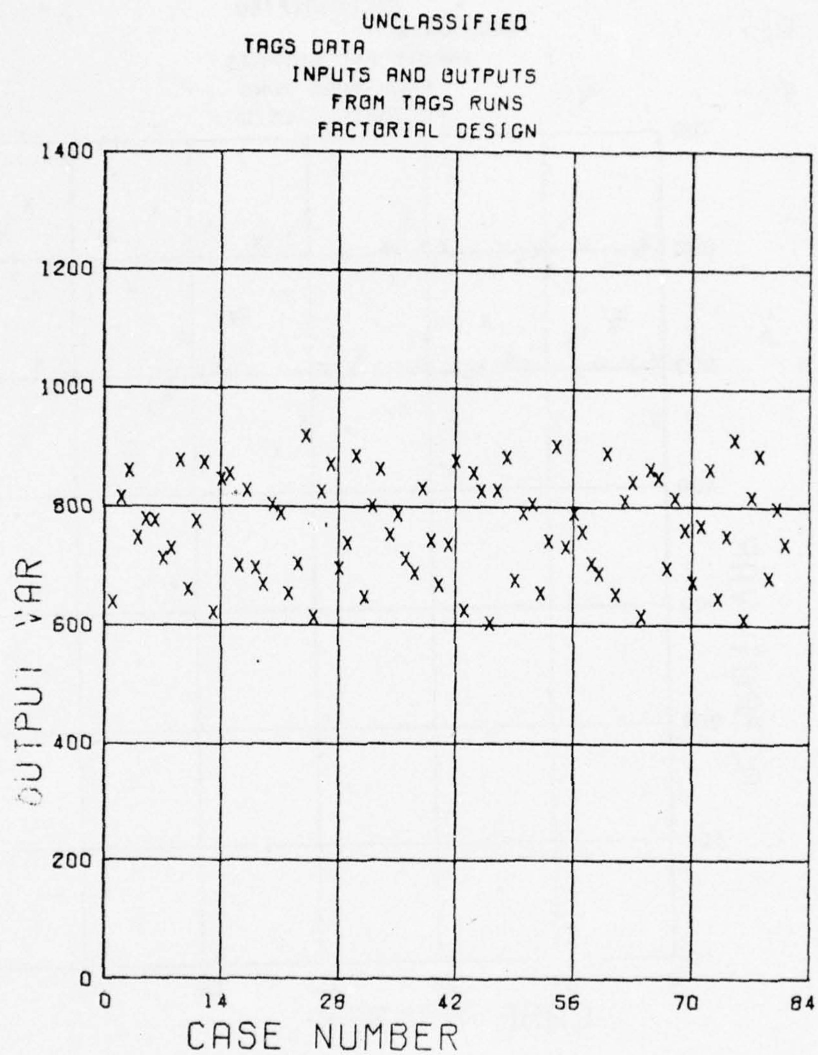


Figure VIII, Blue Fighter Aircraft Losses

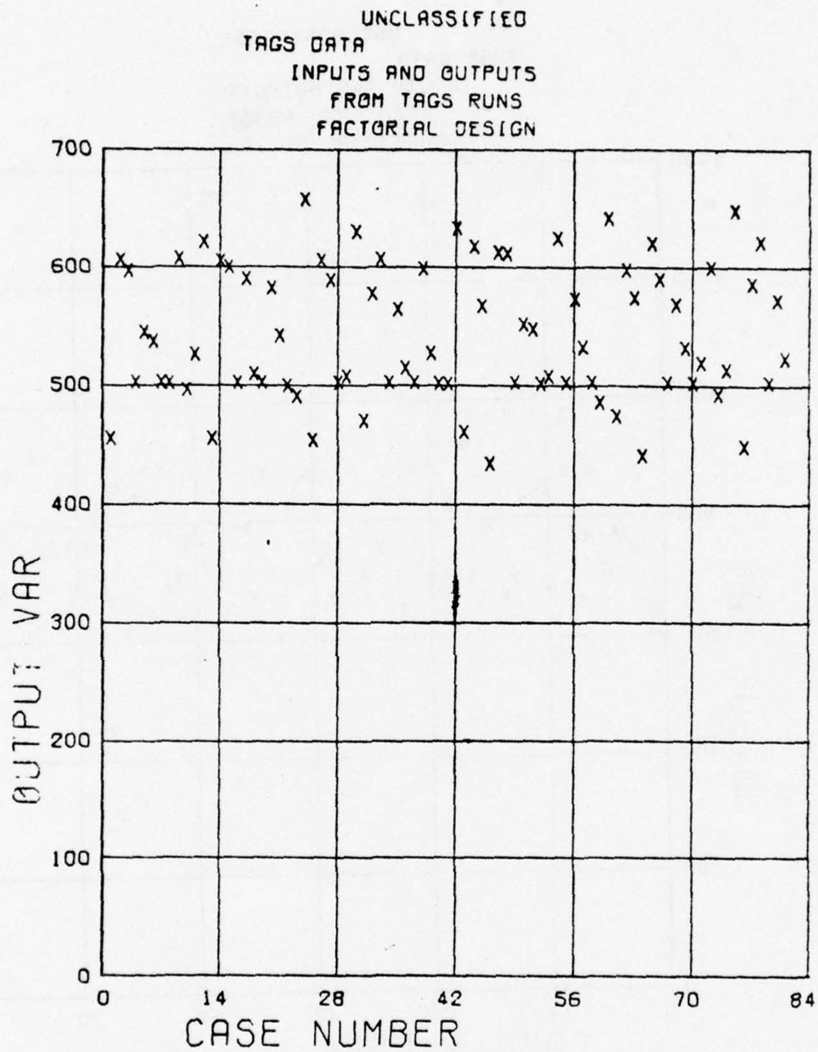


Figure IX, Blue Bomber Aircraft Losses

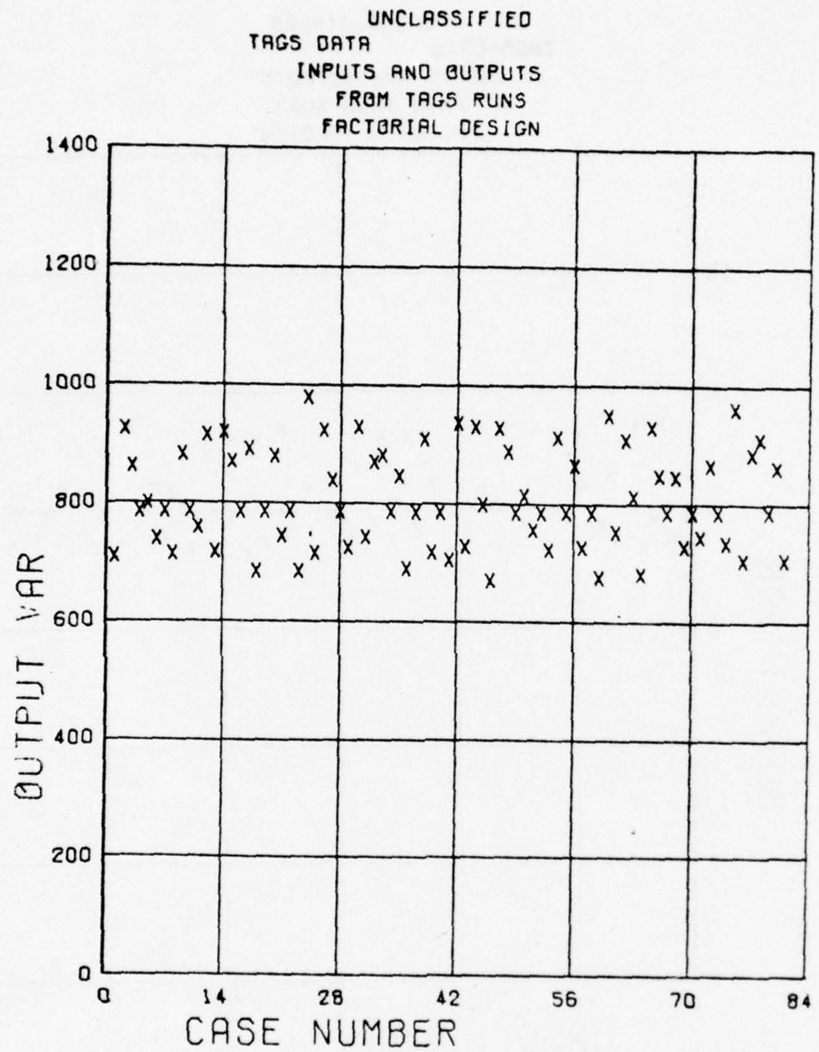


Figure X, Blue Attacker Aircraft Losses

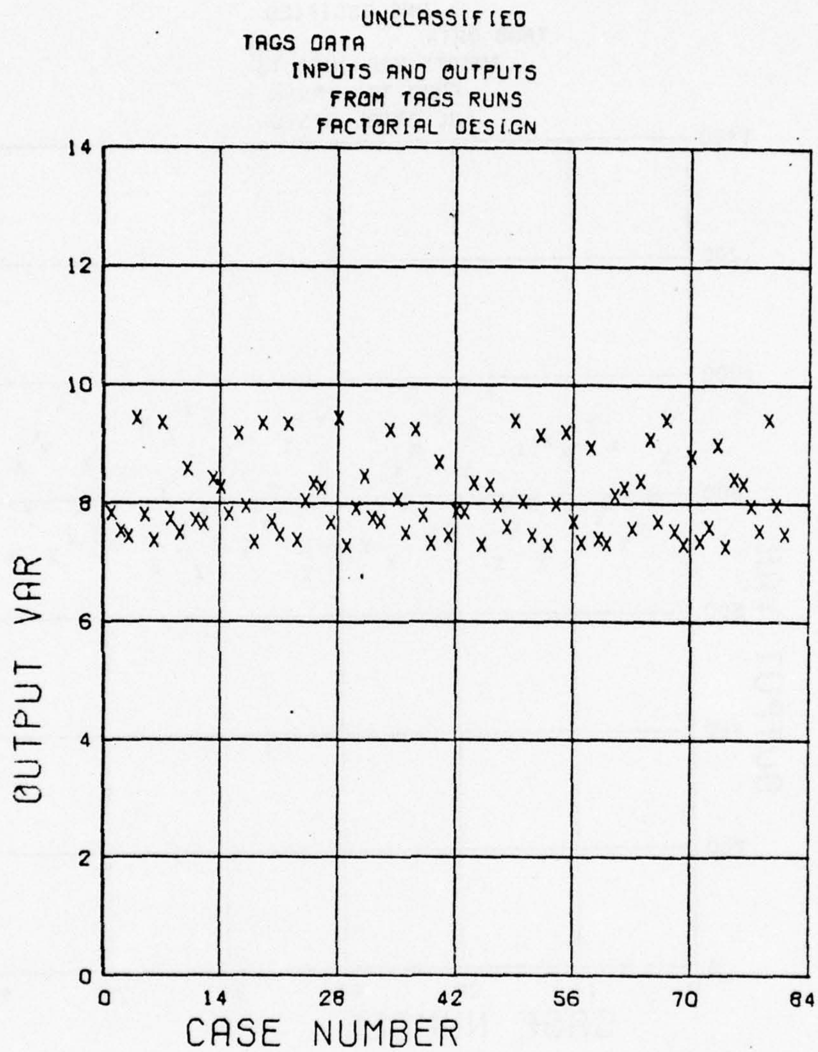


Figure XI, Blue Division Losses to Ground

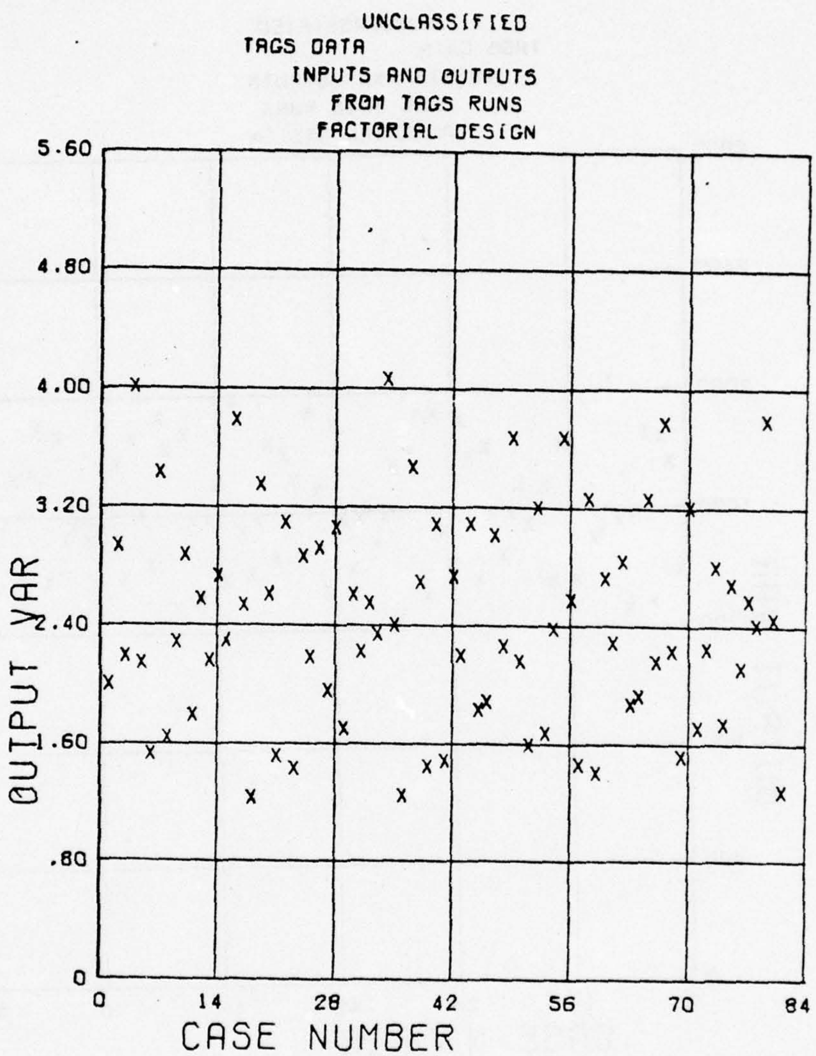


Figure XII, Blue Division Losses to Air Forces

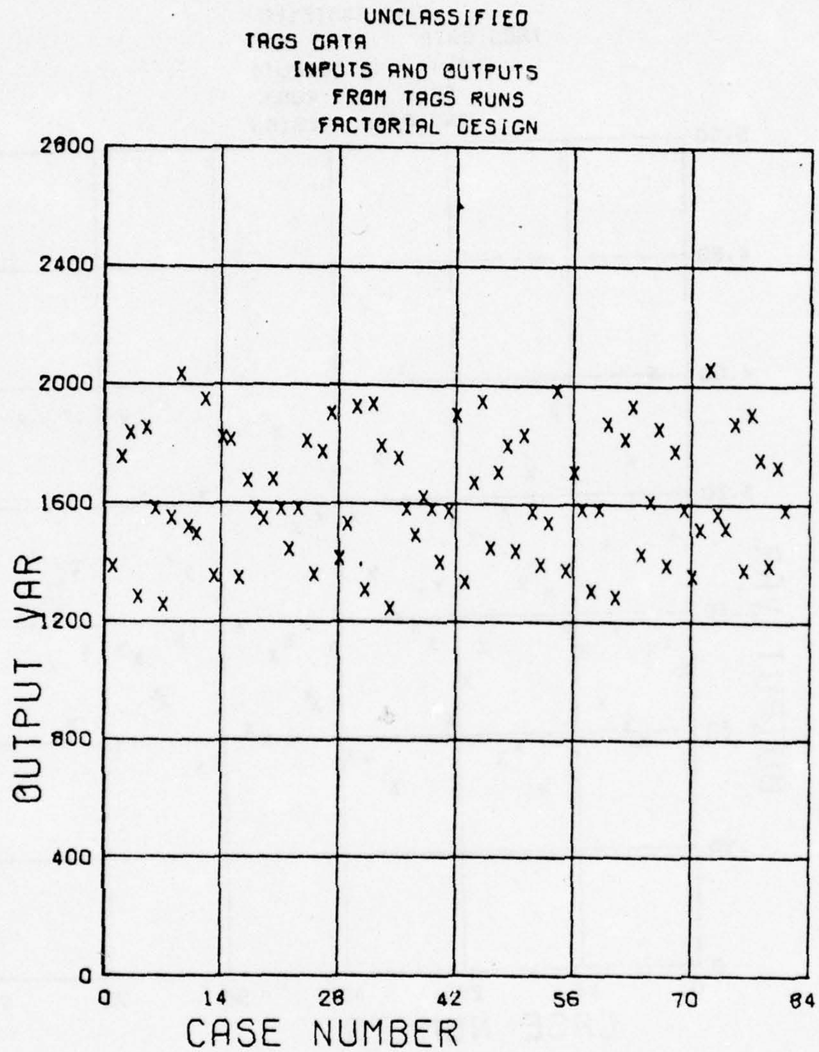


Figure XIII, Red Fighter Aircraft Losses

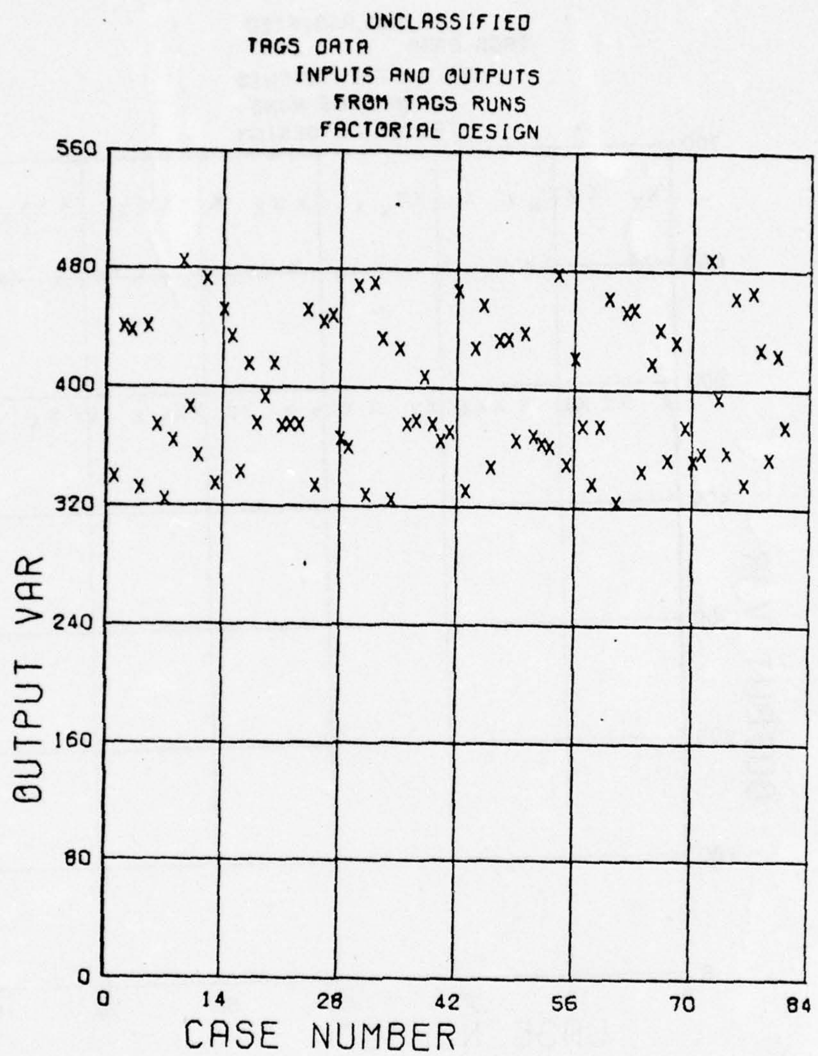


Figure XIV, Red Attacker Aircraft Losses

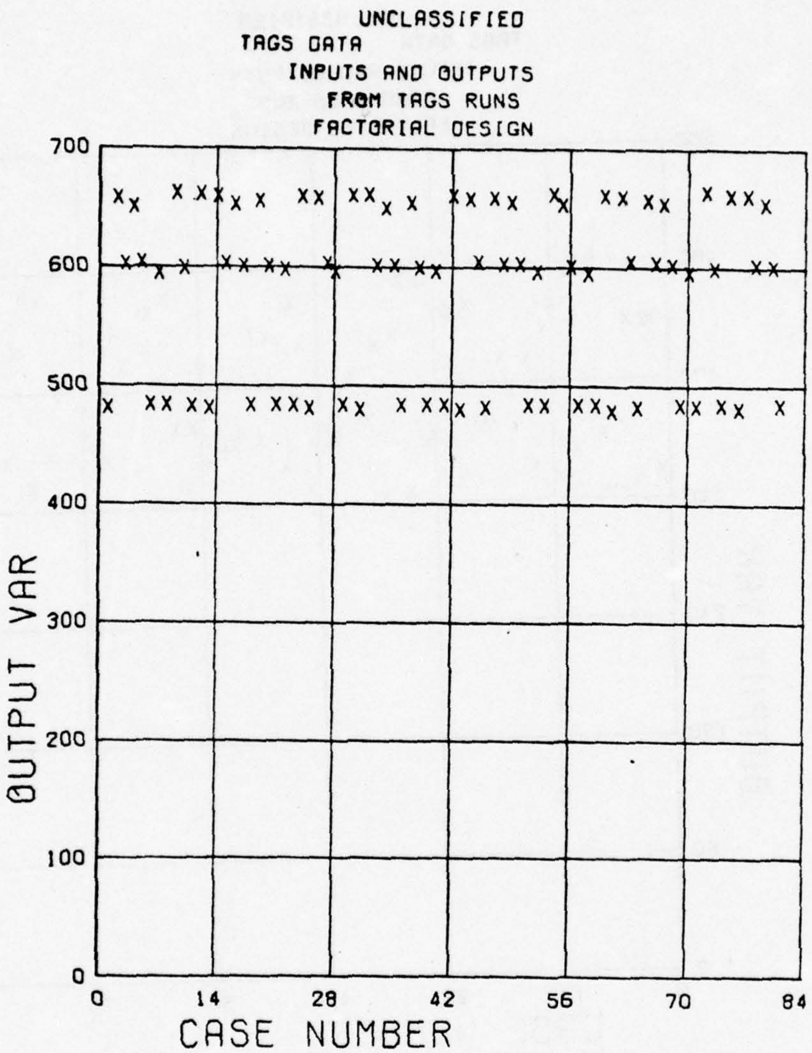


Figure XV, Red Bomber Aircraft Losses

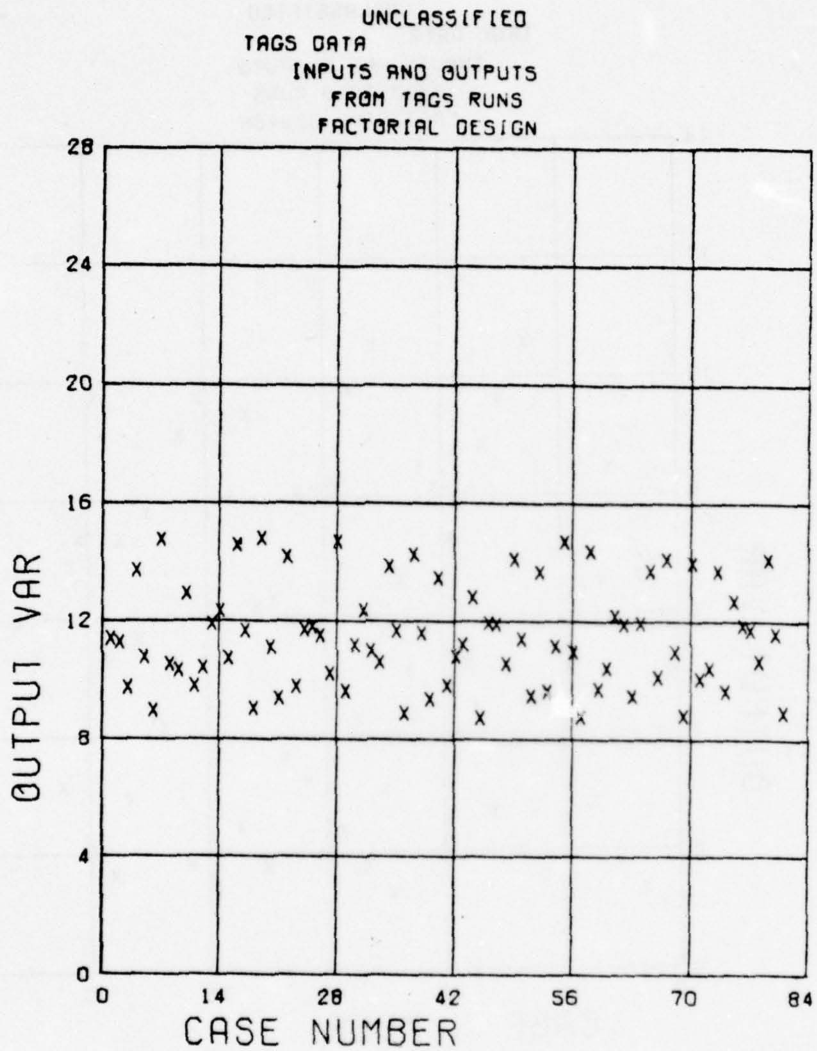


Figure XVI, Red Division Losses to Ground

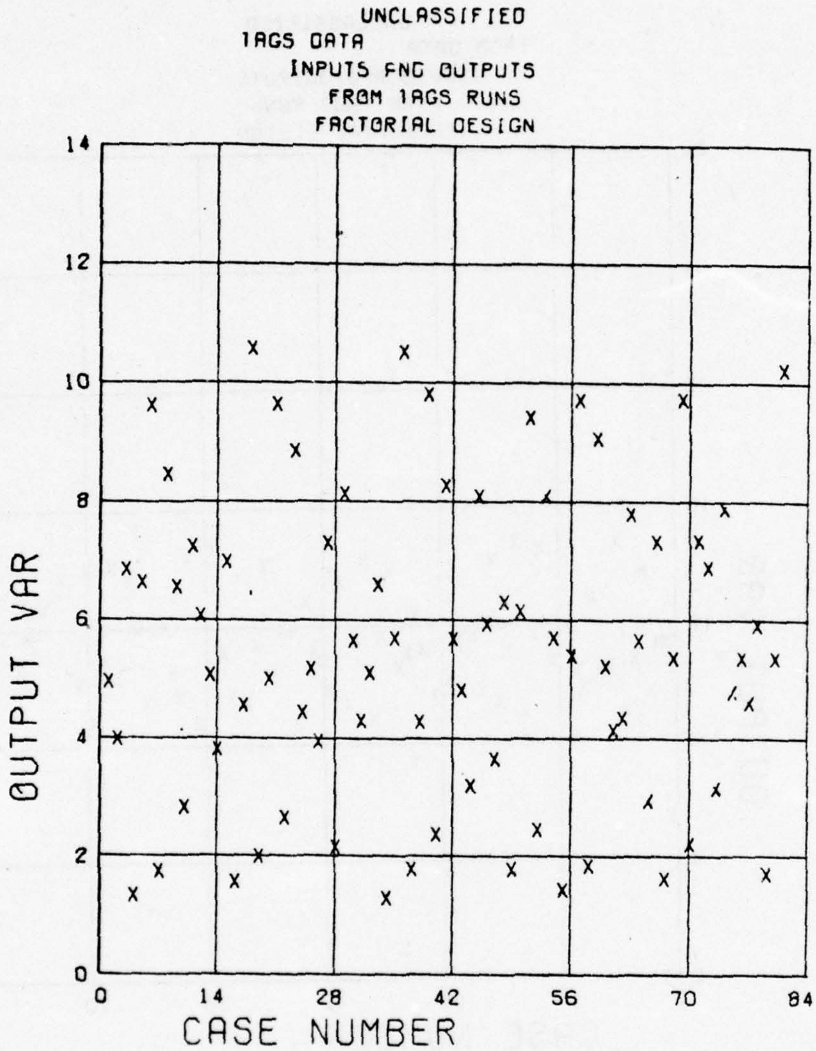


Figure XVII, Red Division Losses to Air Forces

Table V, Means and Standard Deviations for 17 Variables

TEST	MEAN	S.D.
BLUE DIV	19.247	2.482
BLUE TOT A/C	2413.000	311.359
BL A/C SHLT	989.000	128.018
RED DIV	20.773	2.682
RED TOT A/C	2943.333	380.401
RED A/C SHLT	1415.000	182.539
FEBA	-69.046	19.464
BFTR LOSS	762.296	86.725
B ATT LOSS	807.198	82.370
B BMR LOSS	542.951	56.560
RDIV LOSS TOGR	8.083	.686
BDIV LOSS 2AIR	2.443	.710
R FTR LOSS	1618.667	211.980
R ATT LOSS	396.173	47.486
R BMR LOSS	579.580	73.039
RDIV LOSS TOGR	11.467	1.743
RDIV LOSS 2AIR	5.452	2.559

matrix is symmetric, only the upper triangular portion of the matrix is required in order to impart information inherent in the whole matrix. Thus, only 153 correlations need reporting, a savings of 47 percent in storage and reporting costs. In general the savings is (n-1)/(2n) x 100 percent, where n is the number of variables in the data base. Table VI lists the upper triangular portion of the correlation matrix for six input variables and 11 output variables. Usually of greater interest is the examination of those pairs of variables having high correlations. Table VII lists pairs of variables having correlation coefficients whose absolute value exceeds 80 percent.

b. Multiple Regression

(1) Criterion or Dependent Variable. Frequently, a need arises to predict or "explain" variation in values of one variable by considering variation in values of several other variables. The variable whose value is predicted is known as the "criterion" or "dependent" variable; the other variables are known as "predictor" or "independent" variables. A multiple regression equation, an extension of the regression technique used to analyze a pair of variables, is derived by applying the Least Squares Principle. This principle seeks the minimum sum of squared deviations of the predicted value of the criterion variable from the actual value of the criterion variable.

(2) Delineators. The program displays the determinant of the correlation matrix, the multiple correlation coefficient, the square of the multiple correlation coefficient, and the beta (net regression) coefficients. The program also displays the F-value and degress of freedom for a test of statistical significance. In addition, for each predictor variable, the program displays:

(a) R-Squared. The square of the regression coefficient.

(b) Correlation. The correlation between the predictor variable and the criterion variable. This is displayed in the column labeled R(CRIT).

(c) Product. The product of the regression coefficient and the predictor-criterion correlation, a measure of the contribution of the predictor variable to the regression equation. This is displayed in the column labeled BETA*R.

(d) Ratio. The ratio of the predictor-criterion correlation coefficient to the multiple correlation coefficient is displayed in the column labeled STRUC R. The program also

Table VI, Upper Triangular Correlation Matrix

C O R R E L A T I O N   M A T R I X

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
ROW 1	1.00	-.00	.00	.00	.00	.00	-.01	-.00	-.01	-.00	-.01	-.00	-.00	.00	.00	.01	-.00
ROW 2		1.00	-.00	.00	-.00	-.00	.75	.81	.39	.74	-.79	-.56	.80	.70	.03	-.79	.73
ROW 3			1.00	-.00	.00	.00	.09	-.22	-.14	-.12	-.09	-.09	.16	.14	.01	-.13	.15
ROW 4				1.00	.00	.00	-.01	.00	-.00	.00	.02	.00	-.00	-.00	.00	.02	-.00
ROW 5					1.00	-.00	-.57	.50	.81	.61	.45	.78	.42	.59	1.00	.52	-.65
ROW 6						1.00	-.10	.12	.22	.16	.04	.23	-.25	-.20	-.01	.07	-.15
ROW 7							1.00	.36	-.05	.27	-.95	-.90	.52	.35	-.48	-.91	.88
ROW 8								1.00	.81	.96	-.43	-.03	.83	.84	.53	-.36	.21
ROW 9									1.00	.91	.01	.43	.66	.80	.82	.14	-.32
ROW 10										1.00	-.33	.08	.93	.89	.63	-.24	.08
ROW 11											1.00	.84	-.54	-.37	.42	.96	-.87
ROW 12												1.00	-.24	-.03	.75	.89	-.95
ROW 13													1.00	.97	.45	-.48	.35
ROW 14														1.00	.62	-.29	.15
ROW 15															1.00	.49	-.62
ROW 16																1.00	-.95
ROW 17																	1.00

PRT•S H.PRTFILE8
FURPUR 27R1 RL71-3 05/18/76 10:47:21

Table VII, High Correlation Variables

RAW	COR.	BETWEEN	BLUE TOT A/C	AND	BFTR LOSS	= .81
RAW	COR.	BETWEEN	RED TOT A/C	AND	B ATT LOSS	= .81
RAW	COR.	BETWEEN	RED TOT A/C	AND	R BMR LOSS	=1.00
RAW	COR.	BETWEEN	FEBA	AND	BDIVLOSSTOGR	=-.95
RAW	COR.	BETWEEN	FEBA	AND	BDIVLOSS2AIR	=-.90
RAW	COR.	BETWEEN	FEBA	AND	RDIVLOSSTOGR	=-.91
RAW	COR.	BETWEEN	FEBA	AND	RDIVLOSS2AIR	= .88
RAW	COR.	BETWEEN	BFTR LOSS	AND	B ATT LOSS	= .81
RAW	COR.	BETWEEN	BFTR LOSS	AND	B BMR LOSS	= .96
RAW	COR.	BETWEEN	BFTR LOSS	AND	R FTR LOSS	= .83
RAW	COR.	BETWEEN	BFTR LOSS	AND	R ATT LOSS	= .84
RAW	COR.	BETWEEN	B ATT LOSS	AND	B BMR LOSS	= .91
RAW	COR.	BETWEEN	B ATT LOSS	AND	R BMR LOSS	= .82
RAW	COR.	BETWEEN	B BMR LOSS	AND	R FTR LOSS	= .83
RAW	COR.	BETWEEN	B BMR LOSS	AND	R ATT LOSS	= .89
RAW	COR.	BETWEEN	BDIVLOSSTOGR	AND	RDIVLOSS2AIR	= .84
RAW	COR.	BETWEEN	BDIVLOSSTOGR	AND	RDIVLOSSTOGR	= .96
RAW	COR.	BETWEEN	BDIVLOSSTOGR	AND	RDIVLOSS2AIR	=-.87
RAW	COR.	BETWEEN	BDIVLOSS2AIR	AND	RDIVLOSSTOGR	= .89
RAW	COR.	BETWEEN	BDIVLOSS2AIR	AND	RDIVLOSS2AIR	=-.95
RAW	COR.	BETWEEN	R FTR LOSS	AND	R ATT LOSS	= .97
RAW	COR.	BETWEEN	RDIVLOSSTOGR	AND	RDIVLOSS2AIR	=-.95

displays beta coefficients as they are for standardized data. The standardized beta weights are considered adjustments for the fact that the predictors and criterion variables may have come from different distributions. Thus, the standardized beta weights give the user another measure of the relative contributions of the predictor variables to the regression equation. These weights are displayed under the heading of WEIGHTS.

(3) Demonstrability. To demonstrate capability of the Multivariate Analysis System in this area, the decision was made to regress each output variable against the six input variables. Thus, there are 11 multiple regressions - one for each output variable. Results of the multiple regressions are displayed in Tables VIII - XVIII.

(4) Explanation. For explanatory purposes, Table VIII is summarized below:

Regression equation

FEBA = -98.146 + (-.011) (Blue Divisions) + (.745) (Blue Aircraft)
+ (.085) (Blue Aircraft Shelters) + (-.01) (Red Divisions)
+ (-.512) (Red Aircraft) + (-.104) (Red Aircraft Shelters)

Multiple R = 0.914

Multiple R^2 = 0.836, meaning that 83.6 percent of the variance in the original data is preserved (or accounted for) by the regression equation.

c. Principal Components

(1) Parameters. Analysis involving intercorrelations among variables must contend with parameters for the P means, the P variances, plus the correlations. All together there are (P+3) P/2 of the parameters to be estimated and interpreted. For even moderately sized P, this can be a large number. For example, if P = 15, there are 135 parameters. If P = 25, there are 350 parameters to estimate and interpret. In the special case that the variables are uncorrelated, there are only 2P parameters (P means and P variances). In the example above if P = 25, there are now only 50 parameters (instead of 350) to be estimated and interpreted. Moreover, a host of mathematical and statistical techniques become much more simpler in the uncorrelated case.

Table VIII, Multiple Regression for Dependent Variable FEBA

```
MULTIPLE CORRELATION FOR    7 VARIABLES AND        81 OBSERVATIONS
DEPENDENT     VARIABLE IS FEBA
DETERMINANT =            .10009+01
MULTIPLE R-SQUARE =          .836
MULTIPLE R =           .914
F. FOR ANALYSIS OF VARIANCE ON R =          62.861
N.D.F.1 =   6.         N.D.F.2 =              74.
PREDICTOR       BETA   BETASQ          R(CRIT)   BETA*R   STRUC  R
BLUE DIV       -.011    .000            -.011     .000     -.012
BLUE TOT A/C    .745    .555             .745     .555      .815
BL A/C SHLT     .085    .007             .085     .007      .093
RED DIV        -.010    .000            -.010     .000     -.011
RED TOT A/C    -.512    .262            -.512     .262     -.560
RED A/C SHLT   -.104    .011            -.104     .011     -.114
WEIGHTS
BLUE DIV              -.085
BLUE TOT A/C           .047
BL A/C SHLT            .013
RED DIV               -.076
RED TOT A/C           -.026
RED A/C SHLT          -.011
INTERCEPT CONSTANT =     -98.146
```

Table IX, Multiple Regression for Dependent Variable BFTRLOSS

MULTIPLE CORRELATION FOR 7 VARIABLES AND 81 OBSERVATIONS
DEPENDENT VARIABLE IS BFTR LOSS

DETERMINANT = .10000+01

MULTIPLE R-SQUARE = .968

MULTIPLE R = .984

F• FOR ANALYSIS OF VARIANCE ON R = 372.006

N.D.F.1 = 6. N.D.F.2 = 74.

PREDICTOR	BETA	BETASQ	R(CRIT)	BETA*R	STRUC R
BLUE DIV	-.000	.000	-.000	.000	-.000
BLUE TOT A/C	.810	.656	.810	.655	.823
BL A/C SHLT	-.219	.048	-.219	.048	-.223
RED DIV	.003	.000	.003	.000	.003
RED TOT A/C	.500	.250	.500	.250	.509
RED A/C SHLT	.115	.013	.115	.013	.117

WEIGHTS	
BLUE DIV	-.008
BLUE TOT A/C	.226
BL A/C SHLT	-.149
RED DIV	.081
RED TOT A/C	.114
RED A/C SHLT	.055

INTERCEPT CONSTANT = -50.292

Table X, Multiple Regression for Dependent Variable BATTLOSS

MULTIPLE CORRELATION FOR 7 VARIABLES AND 81 OBSERVATIONS
DEPENDENT VARIABLE IS BATTLOSS
DETERMINANT = .1000D+01
MULTIPLE R-SQUARE = .870
MULTIPLE R = .933
F. FOR ANALYSIS OF VARIANCE ON R = 87.665
N.D.F.1 = 6. N.D.F.2 = 74.

PREDICTOR	BETA	BETASQ	R(CRIT)	BETA.R	STRUC R
BLUE DIV	-.006	.000	-.006	.000	-.007
BLUE TOT A/C	.392	.154	.392	.154	.420
BL A/C SHLT	-.141	.020	-.141	.020	-.151
RED DIV	-.001	.000	-.001	.000	-.001
RED TOT A/C	.806	.649	.806	.649	.864
RED A/C SHLT	.219	.048	.219	.048	.234

WEIGHTS	
BLUE DIV	-.202
BLUE TOT A/C	.104
BL A/C SHLT	-.091
RED DIV	-.032
RED TOT A/C	.174
RED A/C SHLT	.099

INTERCEPT CONSTANT = -2.460

Table XI, Multiple Regression for Dependent Variable BBMRLOSS

```
MULTIPLE CORRELATION FOR    7 VARIABLES AND      81 OBSERVATIONS
DEPENDENT     VARIABLE IS R BMR LOSS
DETERMINANT =              .10000+01
MULTIPLE R-SQUARE =           .953
MULTIPLE R =          .976
F. FOR ANALYSIS OF VARIANCE ON R =       248.178
N.D.F.1 =  6.          N.D.F.2 =          74.
PREDICTOR       BETA   BETASQ           R(CRIT)   BETA*R   STRUC  R
BLUE DIV       -.002    .000             -.002     .000    -.002
BLUE TOT A/C   .738     .545              .738     .545     .756
BL A/C SHLT   -.123     .015             -.123     .015    -.126
RED DIV        .000     .000              .000     .000     .000
RED TOT A/C    .607     .368              .607     .368     .621
RED A/C SHLT   .157     .025              .157     .025     .161
WEIGHTS
BLUE DIV             -.041
BLUE TOT A/C          .134
BL A/C SHLT          -.054
RED DIV               .010
RED TOT A/C           .090
RED A/C SHLT          .049
INTERCEPT CONSTANT =       -60.903
```

Table XII, Multiple Regression for Dependent Variable RFTRLOSS

```
DEPENDENT    VARIABLE IS R FTR LOSS
DETERMINANT =          .10000+01
MULTIPLE R-SQUARE =         .896
MULTIPLE R =        .946
F. FOR ANALYSIS OF VARIANCE ON R =      105.792
N.D.F.1 =  6.       N.D.F.2 =          74.
PREDICTOR        BETA   BETASQ      R(CRIT)   BETA*R   STRUC  R
BLUE DIV        -.002    .000        -.002     .000     -.002
BLUE TOT A/C     .797    .636         .797     .636      .843
BL A/C SHLT      .165    .027         .165     .027      .174
RED DIV         -.001    .000        -.001     .000     -.001
RED TOT A/C      .415    .172         .415     .172      .439
RED A/C SHLT    -.246    .060        -.246     .060     -.260
WEIGHTS
BLUE DIV              -.187
BLUE TOT A/C           .543
BL A/C SHLT            .272
RED DIV               -.089
RED TOT A/C            .231
RED A/C SHLT          -.286
INTERCEPT CONSTANT =   -233.138
```

Table XIII, Multiple Regression for Dependent Variable RATTLOSS

MULTIPLE CORRELATION FOR 7 VARIABLES AND 81 OBSERVATIONS
DEPENDENT VARIABLE IS R ATT LOSS
DETERMINANT = .10000+01
MULTIPLE R-SQUARE = .887
MULTIPLE R = .942
F. FOR ANALYSIS OF VARIANCE ON R = 96.661
N.D.F.1 = 6. N.D.F.2 = 74.

PREDICTOR	BETA	BETASQ	R(CRIT)	BETA*R	STRUC R
BLUE DIV	.000	.000	.000	.000	.000
BLUE TOT A/C	.695	.483	.695	.483	.738
BL A/C SHLT	.137	.019	.137	.019	.145
RED DIV	-.003	.000	-.003	.000	-.003
RED TOT A/C	.588	.345	.588	.345	.624
RED A/C SHLT	-.198	.039	-.198	.039	-.210

WEIGHTS	
BLUE DIV	.007
BLUE TOT A/C	.106
BL A/C SHLT	.051
RED DIV	-.053
RED TOT A/C	.073
RED A/C SHLT	-.052

INTERCEPT CONSTANT = -52.275

Table XIV, Multiple Regression for Dependent Variable BDIVLOSSTOGR

```
MULTIPLE CORRELATION FOR    7 VARIABLES AND        81 OBSERVATIONS
DEPENDENT    VARIABLE IS BDIVLOSSTOGR
DETERMINANT =          .1000D+01
MULTIPLE R-SQUARE =         .839
MULTIPLE R =       .916
F: FOR ANALYSIS OF VARIANCE ON R =       64.06[illegible]
N.D.F.1 =  6.        N.D.F.2 =        74.
PREDICTOR        BETA   BETASQ         R(CRIT)   BETA*R   STDUC R
BLUE DIV        -.011    .000          -.011     .000     -.01[illegible]
BLUE TOT A/C    -.793    .629          -.793     .629     -.866
BL A/C SHLT     -.087    .008          -.087     .008     -.095
RED DIV          .022    .000           .022     .000      .024
RED TOT A/C      .447    .200           .447     .200      .488
RED A/C SHLT     .036    .001           .036     .001      .040
WEIGHTS
BLUE DIV               -.007
BLUE TOT A/C           -.002
BL A/C SHLT            -.000
RED DIV                 .006
RED TOT A/C             .001
RED A/C SHLT            .000
INTERCEPT CONSTANT =     10.132
```

Table XV, Multiple Regression for Dependent Variable BDIVLOSS2AIR

MULTIPLE CORRELATION FOR 7 VARIABLES AND 81 OBSERVATIONS

DEPENDENT VARIABLE IS BDIVLOSS2AIR

DETERMINANT = .10000+01

MULTIPLE R-SQUARE = .978

MULTIPLE R = .989

F. FOR ANALYSIS OF VARIANCE ON R = 555.850

N.D.F.1 = 6. N.D.F.2 = 74.

PREDICTOR	BETA	BETASQ	R(CRIT)	BETA*R	STRUC R
BLUE DIV	-.001	.000	-.001	.000	-.001
BLUE TOT A/C	-.557	.310	-.557	.310	-.563
BL A/C SHLT	-.092	.009	-.092	.009	-.093
RED DIV	.001	.000	.001	.000	.001
RED TOT A/C	.780	.608	.780	.608	.788
RED A/C SHLT	.227	.052	.227	.052	.230

WEIGHTS	
BLUE DIV	-.000
BLUE TOT A/C	-.001
BL A/C SHLT	-.001
RED DIV	.000
RED TOT A/C	.001
RED A/C SHLT	.001

INTERCEPT CONSTANT = .475

Table XVI, Multiple Regression for Dependent Variable
RBMRLOSS

MULTIPLE CORRELATION FOR 7 VARIABLES AND 81 OBSERVATIONS
DEPENDENT VARIABLE IS R BMR LOSS *
DETERMINANT = .10000+01
MULTIPLE R-SQUARE = 1.000 *
MULTIPLE R = 1.000
F. FOR ANALYSIS OF VARIANCE ON R = 48594.431
N.D.F.1 = 6. N.D.F.2 = 74.

PREDICTOR	BETA	BETASQ	R(CRIT)	BETA*R	STRUC R
BLUE DIV	.000	.000	.000	.000	.000
BLUE TOT A/C	.035	.001	.035	.001	.035
BL A/C SHLT	.009	.000	.009	.000	.009
RED DIV	.001	.000	.001	.000	.001
RED TOT A/C	~~.999~~	.998	.999	.998	.999
RED A/C SHLT	-.010	.000	-.010	.000	-.010

WEIGHTS

BLUE DIV	.003
BLUE TOT A/C	.008
BL A/C SHLT	.005
RED DIV	.024
RED TOT A/C	.192
RED A/C SHLT	-.004

INTERCEPT CONSTANT = -5.237

Table XVII, Multiple Regression for Dependent Variable RDIVLOSSTOGR

MULTIPLE CORRELATION FOR 7 VARIABLES AND 81 OBSERVATIONS
DEPENDENT VARIABLE IS RDIVLOSSTOGR
DETERMINANT = .10000+01
MULTIPLE R-SQUARE = .923
MULTIPLE R = .961
F FOR ANALYSIS OF VARIANCE ON R = 147.841
N.D.F.1 = 6. N.D.F.2 = 74.

PREDICTOR	BETA	BETASQ	R(CRIT)	BETA*R	STRUC R
BLUE DIV	.011	.000	.011	.000	.011
BLUE TOT A/C	-.792	.627	-.792	.627	-.824
BL A/C SHLT	-.130	.017	-.130	.017	-.136
RED DIV	.018	.000	.018	.000	.019
RED TOT A/C	.523	.274	.523	.274	.545
RED A/C SHLT	.070	.005	.070	.005	.072

WEIGHTS	
BLUE DIV	.008
BLUE TOT A/C	-.004
BL A/C SHLT	-.002
RED DIV	.012
RED TOT A/C	.002
RED A/C SHLT	.001

INTERCEPT CONSTANT = 15.516

Table XVIII, Multiple Regression for Dependent Variable RDIVLOSSTOAI

```
@XQT X.ABS8
MULTIPLE CORRELATION FOR    7 VARIABLES AND       81 OBSERVATIONS
DEPENDENT   VARIABLE IS RDIVLOSS2AIR
DETERMINANT =          .10000+01
MULTIPLE R-SQUARE =       .988
MULTIPLE R =       .994
F* FOR ANALYSIS OF VARIANCE ON R =     1048.644
N.D.F.1 =  6.      N.D.F.2 =        74.
PREDICTOR       BETA   BETASQ         R(CRIT)   BETA*R   STRUC  R
BLUE DIV       -.003     .000          -.003     .000     -.004
BLUE TOT A/C   .726      .527           .726     .527      .730
BL A/C SHLT    .145      .021           .145     .021      .146
RED DIV       -.003      .000          -.003     .000     -.003
RED TOT A/C  -.647       .418          -.647     .418     -.650
RED A/C SHLT-.147        .022          -.147     .022     -.148
WEIGHTS
BLUE DIV              -.004
BLUE TOT A/C           .006
BL A/C SHLT            .003
RED DIV               -.003

RED TOT A/C           -.004
RED A/C SHLT          -.002
INTERCEPT CONSTANT =       4.055
```

(2) Data Transformation. Generally, original data is correlated. The correlation can, however, be removed by applying a suitable transformation. This is precisely what principal components does. The first principal component is that normalized linear combination (i.e., sum of squares of coefficients equals one) of the original variables having maximum variance. The second principal component is constructed as a dimension uncorrelated with the first principal component which has maximum variance. The process continues until there are no more uncorrelated dimensions to construct. Geometrically, this is equivalent to finding the line so that the sum of squared distances is minimized from the data points to the line.

(3) Virtues. Principal components may have the virtues of parsimony, orthogonality, increased reliability and increased normality over the observation measures from which they are derived. Principal components reveal not only how several measures of a domain can be combined to produce maximum discrimination among individuals along a single dimension, but also that several dimensions are required to define the domain under investigation adequately.

(4) Factors. Besides generating a full-rank set of principal component factors, the Multivariate Analysis System provides eigenvalues and eigenvectors for the correlation matrix as well as the fractions of the total variance accounted for by each factor. Also available are the factor score coefficients for each variable, ghe inverse of the correlation matrix, and the factor pattern matrix. Since all work by the Multivariate Analysis System is for a set of orthogonal factors, the factor pattern matrix is the same as the factor structure matrix, each of which is the matrix of correlations between the set of factors and the set of selected variables.

(5) Coordinate Transformation. Subsequent to determining the principal components, the analyst may want to recast the problem at hand into a different coordinate system. The reason for this is that the coordinates of a point depend on the location of that point relative to the origin of the coordinate axes. Therefore, changing the location of the origin alters the coordinates of the point. The usual transformation of coordinate axes uses is a "rotation" or "orthogonal transformation". The object is to simplify the principal components to make them more easily interpretable.

(6) Comparison. The quartimax method maximizes variance of squared factor loadings, while the varimax simplifies columns, rather than rows, of the factor pattern matrix. In general, the varimax method makes it earier for the user to name the resulting factors; the quartimax method permits the user to understand the impact of the variables upon the factors.

(7) Incompressibility. Table XIX shows a principal components program done on the six input variables. The fact is known, from the original experimental design, that original inputs are formed on uncorrelated groups of variables. Thus, a principal components exercise should not develop any new information. The results in Table XIX confirm this. The fact that each factor has an associated eigenvalue equal to one indicates that each factor contains exactly the same percentage of information as the original variables. Additionally, the factor pattern matrix contains high values in each column. This indicates again that the information contained in the original six variables cannot be compressed into fewer than six dimensions without a serious loss of information. This is further illustrated by the "sphericity" test, a test of the null hypothesis that the original variables are uncorrelated. The test statistic is chi-square distributed. The calculated test statistic is 0.00; therefore, the null hypothesis cannot be rejected. Thus, further credence is lent to the view that the information is incompressible.

(8) Exercise Results. Following this use of the principal components routine on the six input variables is a similar exercise on the 11 output variables. Means and standard deviations are followed by an application of the method of principal components and rotation of axes. Tables XX-XXIV display the results of this exercise. The high percentage (94.0) of variance is accounted for only by the first two principal components in Table XXII. Additionally, after varimax rotation, the first two factors admit to improved interpretability without sacrificing any of the information inherent in them. The factors in Table XXIV have been labeled as Ground Losses and Aircraft Losses, respectively.

d. Canonical Correlation

(1) Background. Canonical Correlation is a technique for analyzing and describing the relationships between two sets of random variables wherein each set consists of multiple measures.

Table XIX, Principal Components

PRINCIPAL COMPONENTS PROGRAM.
6 TESTS 81 SUBJECTS 0 CRITERIA
DETERMINANT OF R = .1000000+01
SPHERICITY TEST, CHI-SQUARE= .00 NDCF= 15

FACTOR	EIGENVALUE	PERCENT TRACE	CUM.PERCENT	N.D.F.	CHI-SQUARE
1	1.0000	16.7	16.7	15	.00
2	1.0000	16.7	33.3	10	.00
3	1.0000	16.7	50.0	6	.00
4	1.0000	16.7	66.7	3	.00
5	1.0000	16.7	83.3	1	.00
6	1.0000	16.7	100.0	0	.00

FACTOR PATTERN. FACTORS ARE COLUMNS, TESTS ARE ROWS.

BLUE DIV	.664	.127	.127	-.004	-.139	-.674
BLUE TOT A/C	-.034	.214	.214	.314	.339	-.178
BL A/C SHLT	-.014	.314	.314	.813	-.162	-.043
RED DIV	.626	-.055	-.055	.103	-.348	.667
RED TOT A/C	.358	-.517	-.517	.207	.751	.116
RED A/C SHLT	.194	.754	.754	-.432	.393	.232

Table XX, Means and Standard Deviations for 11 Output Variables

TEST	MEAN	S.D.
FEBA	-64.046	19.464
BFTR LOSS	742.296	86.725
B ATT LOSS	807.198	82.370
B BMR LOSS	542.951	56.560
BDIVLOSSTOGR	8.083	.686
BDIVLOSS2AIR	2.443	.710
R FTR LOSS	1618.667	211.980
R ATT LOSS	396.173	47.486
R BMR LOSS	579.580	73.039
RDIVLOSSTOGR	11.467	1.743
RDIVLOSS2AIR	5.452	2.559

Table XXI, High Valued Correlations for 11 Output Variables

RAW COR. BETWEEN	FEBA	AND BDIVLOSSTOGR	=-.95
RAW COR. BETWEEN	FEBA	AND RDIVLOSSTOGR	=-.91
RAW COR. BETWEEN	BFTR LOSS	AND B BMR LOSS	= .96
RAW COR. BETWEEN	B ATT LOSS	AND B BMR LOSS	= .91
RAW COR. BETWEEN	BDIVLOSSTOGR	AND RDIVLOSSTOGR	= .96
RAW COR. BETWEEN	BDIVLOSS2AIR	AND RDIVLOSS2AIR	=-.95
RAW COR. BETWEEN	R FTR LOSS	AND R ATT LOSS	= .97
RAW COR. BETWEEN	RDIVLOSSTOGR	AND RDIVLOSS2AIR	=-.95

Table XXII, Principal Components for 11 Output Variables

PRINCIPAL COMPONENTS PROGRAM.
11 TESTS 81 SUBJECTS 0 CRITERIA
DETERMINANT OF R = $\boxed{.2741581-14}$

FACTOR	EIGENVALUE	PERCENT TRACE	CUM. PERCENT
1	5.6166	51.1	51.1
2	4.7243	42.9	94.0
3	.3101	2.8	96.8
4	.1763	1.6	98.4
5	.0980	.9	99.3
6	.0359	.3	99.6
7	.0213	.2	99.8
8	.0123	.1	100.0
9	.0037	.0	100.0
10	.0010	.0	100.0
11	.0006	.0	100.0

Table XXIII, Factor Pattern Matrix: Correlations Between Principal Components and Underlying Variables

FACTOR PATTERN. FACTORS ARE COLUMNS, TESTS ARE ROWS.

FERA	.883	-.399	-.004	-.221	.029	.070	.047	.046	-.008	.001	.002
BFTR LOSS	.713	.626	-.247	.146	.054	.110	-.044	.007	.000	.006	-.032
B ATT LOSS	.319	.914	-.160	-.175	.026	-.075	-.015	.018	.016	-.003	-.013
B AMR LOSS	.645	.733	-.188	.022	.064	-.063	.041	-.027	-.015	-.006	.013
BDIVLOSSTOGR	-.905	.349	.092	.092	.193	-.021	.004	.063	-.011	-.000	.001
BDIVLOSS2AIR	-.686	.709	-.080	.038	-.130	.006	.045	.033	.029	.009	.008
R FTR LOSS	.819	.494	.285	.018	.012	.025	-.033	.013	.026	-.015	.006
R ATT LOSS	.696	.665	.266	-.036	.018	-.032	-.012	-.014	-.009	.022	.001
R AMR LOSS	-.082	.974	.136	.094	-.105	.035	.058	-.003	-.024	-.008	-.010
RDIVLOSSTOGR	-.869	.447	.047	-.118	.135	.069	.035	-.060	.016	.080	-.001
RDIVLOSS2AIR	.776	-.586	.018	.202	.075	-.035	.072	-.008	.024	.003	-.007

Table XXIV, Rotation of Principal Components

```
ROTATE ORTHOGONAL FACTORS
ENTER NUMBER OF FACTORS TO BE ROTATED, 12
ENTER 0 FOR VARIMAX OR 1 FOR QUARTIMAX
VARIMAX ROTATION OF  2 FACTORS BASED ON 11 TESTS

NEW FACTOR PATTERN. COLUMNS ARE FACTORS
FEBA              .947       .203
BFTR LOSS         .202       .927
B ATT LOSS       -.286       .925
B BMR LOSS        .083       .973
BDIVLOSSTOGR     -.936      -.256
BDIVLOSS2AIR     -.972       .162
R FTR LOSS        .366       .884
R ATT LOSS        .165       .948
R BMR LOSS       -.645       .735
RDIVLOSSTOGR     -.965      -.156
RDIVLOSS2AIR      .973      -.010
TYPE 1 TO OBTAIN SORTED FACTOR PRINTOUT, BY FACTOR.
SORTED FACTOR PATTERN FOR FACTOR   1.
BDIVLOSS2AIR     -.972
RDIVLOSSTOGR     -.965
BDIVLOSSTOGR     -.936
R BMR LOSS       -.645
B ATT LOSS       -.286
B BMR LOSS        .083          Ground Losses
R ATT LOSS        .165
BFTR LOSS         .202
R FTR LOSS        .366
FEBA              .947
RDIVLOSS2AIR      .973
SORTED FACTOR PATTERN FOR FACTOR   2.
BDIVLOSSTOGR     -.256
RDIVLOSSTOGR     -.156
RDIVLOSS2AIR     -.010
BDIVLOSS2AIR      .162
FEBA              .203
R BMR LOSS        .735
R FTR LOSS        .884
B ATT LOSS        .925          Aircraft Losses
BFTR LOSS         .927
R ATT LOSS        .948
B BMR LOSS        .973
PROPORTION OF VARIANCE ACCOUNTED FOR BY EACH FACTOR.
FACTOR   1   VAR. PROP.=   .482
FACTOR   2   VAR. PROP.=   .458
```

Let the first set of random variables consist of p variables, say $X_1, X_2, \ldots, X_p$, while the second set consists of q variables, say, $Y_1, Y_2, \ldots, Y_q$. The objective is then to seek a standardized linear combination of the X's and a standardized linear combination of the Y's so that these two linear compounds have as large a simple correlation coefficient between them as possible.

(2) Coefficients. The geometric view of canonical correlation is that of finding a transformation, or rotation, of the first p coordinate axes and a rotation of the last q coordinate axes to a new p+q-dimensional system which more clearly displays the intercorrelations between the X and Y variables. The procedure of finding maximally correlated linear compounds may be continued, subject to the constraint that subsequent compounds be uncorrelated with previous combinations. In this way, a set of so-called "canonical variables," "canonical weights," and "canonical correlation coefficients" are generated. The purpose for performing this procedure is the hope that, in a multivariate situation, the first few canonical variables will impart a high percentage of the information carried by the entire set of original variables. The level of information, which may be called "a high percentage," is user-defined. It is not unusual to accept levels of 70 percent, 80 percent, or higher. On the other hand, the literature also contains references and examples in which the 50 percent level is used. Thus, the analyst can reduce the number of variables under study to a more comprehensible dimension.

(3) Predictor. In this study the six input variables were chosen as the X, or predictor, variables. For illustrative and investigative purposes, four distinct sets of Y, or criteria variables, were selected with each set used in tandem with the X variables. The first set of Y variables chosen consists of the three variables BFTRLOSS, BATTLOSS, and BBMRLOSS. The second set of Y variables consists of the three variables RFTRLOSS, RATTLOSS, and RBMRLOSS. The third set of Y variables consists of the two variables BDIVLOSSTOGR and BDIVLOSSTOAIR. The fourth set of Y variables consists of the two variables RDIVLOSSTOGR (Red division losses to ground forces) and RDIVLOSSTOAI (Red division losses to air forces).

(4) Predictions. Tables XXV - XXVIII display the statistical outputs from these exercises of the canonical correlation model. Table XXIX lists the observed and predicted values for case number four, in which the variables are Red's division losses to ground forces and to air forces.

Table XXV, Canonical Correlation, Case 1

```
CANONICAL CORRELATIONS
   6 PREDICTORS ON LEFT     2 CRITERIA ON RIGHT
CANONICAL REGRESSION COEFFICIENTS
 .99024    .76823
PREDICTOR CANNONICAL WEIGHTS
BLUE DIV          -.000      -.025
BLUE TOT A/C       .515      -.834
BL A/C SHLT        .092      -.031
RED DIV            .002       .051
RED TOT A/C       -.814      -.423
RED A/C SHLT      -.251      -.349
CRITERION CANNONICAL WEIGHTS
BDIVLOSSTOGR       .140      1.843
BDIVLOSS2AIR     -1.115     -1.474
FACTOR STRUCTURE OF PREDICTORS, COLUMNS ARE FACTORS
BLUE DIV          -.000      -.025
BLUE TOT A/C       .515      -.834
BL A/C SHLT        .092      -.031
RED DIV            .002       .051
RED TOT A/C       -.814      -.423
RED A/C SHLT      -.251      -.349
FACTOR   VARIANCE EXTRACTED   REDUNDANCY
   1               .167          .163
   2               .167          .098
TOTAL VARIANCE EXTRACTED FROM PREDICTORS =        .333
REDUNDANCY OF PREDICTORS GIVEN CRITERIA        .262
NOTE THAT ALL VALUES ARE PROPORTIONS
FACTOR STRUCTURE FOR CRITERIA, COLUMNS ARE FACTORS
BDIVLOSSTOGR      -.798       .603
BDIVLOSS2AIR      -.997       .076
FACTOR   VARIANCE EXTRACTED   REDUNDANCY
   1               .815          .799
   2               .185          .109
TOTAL VARIANCE EXTRACTED FROM CRITERIA =     1.000
TOTAL REDUNDANCY FOR CRITERIA GIVEN PREDICTORS        .908
NOTE THAT ALL VALUES ARE PROPORTIONS
WILKS LAMBDA FOR TOTAL SET =        .0079640
CHI SQUARE FOR TOTAL =    364.8780212
N.D.F. =     12
```

Table XXVI, Canonical Correlation, Case 2

```
CANONICAL CORRELATIONS
   6 PREDICTORS ON LEFT     3 CRITERIA ON RIGHT
CANONICAL REGRESSION COEFFICIENTS
  .99999   .94201   .14740
PREDICTOR CANNONICAL WEIGHTS
BLUE DIV            -.000     -.005     -.130
BLUE TOT A/C         .002      .931     -.354
BL A/C SHLT         -.001      .200      .349
RED DIV             -.001      .001      .132
RED TOT A/C        -1.000      .002     -.000
RED A/C SHLT        -.001     -.306     -.848
CRITERION CANONICAL WEIGHTS
R FTR LOSS           .045     2.053     7.416
R ATT LOSS           .003    -1.082    -8.680
R BMR LOSS         -1.021     -.214     2.025
FACTOR STRUCTURE OF PREDICTORS, COLUMNS ARE FACTORS
BLUE DIV            -.000     -.005     -.130
BLUE TOT A/C         .002      .931     -.354
BL A/C SHLT         -.001      .200      .349
RED DIV             -.001      .001      .132
RED TOT A/C        -1.000      .002     -.000
RED A/C SHLT        -.001     -.306     -.848
FACTOR    VARIANCE EXTRACTED    REDUNDANCY
    1             .167             .167
    2             .167             .148
    3             .167             .004
TOTAL VARIANCE EXTRACTED FROM PREDICTORS =        .500
REDUNDANCY OF PREDICTORS GIVEN CRITERIA        .318
NOTE THAT ALL VALUES ARE PROPORTIONS
FACTOR STRUCTURE FOR CRITERIA, COLUMNS ARE FACTORS
R FTR LOSS          -.413      .904     -.113
R ATT LOSS          -.586      .782     -.213
R BMR LOSS          -.999      .042     -.006
FACTOR    VARIANCE EXTRACTED    REDUNDANCY
    1             .504             .504
    2             .476             .423
    3             .019             .000
TOTAL VARIANCE EXTRACTED FROM CRITERIA =     1.000
TOTAL REDUNDANCY FOR CRITERIA GIVEN PREDICTORS        .927
NOTE THAT ALL VALUES ARE PROPORTIONS
WILKS LAMBDA FOR TOTAL SET =        .0000030
CHI SQUARE FOR TOTAL =    953.6155319
N.D.F. =      18
```

Table XXVII, Canonical Correlation, Case 3

```
CANONICAL CORRELATIONS
    6 PREDICTORS ON LEFT     3 CRITERIA ON RIGHT
CANONICAL REGRESSION COEFFICIENTS
  .99832    .90140    .49440
PREDICTOR CANNONICAL WEIGHTS
BLUE DIV               .003      -.011      .002
BLUE TOT A/C           .952      -.304     -.036
BL A/C SHLT           -.106      -.222     -.965
RED DIV                .002      -.002      .015
RED TOT A/C            .281       .893     -.209
RED A/C SHLT           .059       .246     -.155
CRITERION CANONICAL WEIGHTS
BFTR LOSS              .230       .768     4.478
B ATT LOSS            -.956      2.392     1.344
B BMR LOSS            1.543     -2.483    -5.651
FACTOR STRUCTURE OF PREDICTORS, COLUMNS ARE FACTORS
BLUE DIV               .003      -.011      .002
BLUE TOT A/C           .952      -.304     -.036
BL A/C SHLT           -.106      -.222     -.965
RED DIV                .002      -.002      .015
RED TOT A/C            .281       .893     -.209
RED A/C SHLT           .059       .246     -.155
FACTOR    VARIANCE EXTRACTED    REDUNDANCY
   1                 .167          .166
   2                 .167          .135
   3                 .167          .041
TOTAL VARIANCE EXTRACTED FROM PREDICTORS =        .500
REDUNDANCY OF PREDICTORS GIVEN CRITERIA         .342
NOTE THAT ALL VALUES ARE PROPORTIONS
FACTOR STRUCTURE FOR CRITERIA, COLUMNS ARE FACTORS
BFTR LOSS              .943       .308      .122
B ATT LOSS             .628       .761     -.163
B BMR LOSS             .897       .425     -.119
FACTOR    VARIANCE EXTRACTED    REDUNDANCY
   1                 .697          .694
   2                 .285          .232
   3                 .019          .005
TOTAL VARIANCE EXTRACTED FROM CRITERIA =        1.000
TOTAL REDUNDANCY FOR CRITERIA GIVEN PREDICTORS          .930
NOTE THAT ALL VALUES ARE PROPORTIONS
WILKS LAMBDA FOR TOTAL SET =          .0004764
CHI SQUARE FOR TOTAL =    573.6976852
N.D.F. =       18
```

Table XXVIII, Canonical Correlation, Case 4

```
CANONICAL CORRELATIONS
   6 PREDICTORS ON LEFT     2 CRITERIA ON RIGHT
CANONICAL REGRESSION COEFFICIENTS  .
  .99432    .49117
PREDICTOR CANNONICAL WEIGHTS
BLUE DIV             .004      .048
BLUE TOT A/C        -.737     -.634
BL A/C SHLT         -.146      .055
RED DIV              .004      .099
RED TOT A/C          .645     -.610
RED A/C SHLT         .144     -.460
CRITERION CANONICAL WEIGHTS
RDIVLOSSTOGR         .063     3.172
RDIVLOSS2AIR        -.940     3.030
FACTOR STRUCTURE OF PREDICTORS, COLUMNS ARE FACTORS
BLUE DIV             .004      .048
BLUE TOT A/C        -.737     -.634
BL A/C SHLT         -.146      .055
RED DIV              .004      .099
RED TOT A/C          .645     -.610
RED A/C SHLT         .144     -.460
FACTOR   VARIANCE EXTRACTED    REDUNDANCY
   1               .167            .165
   2               .167            .040
TOTAL VARIANCE EXTRACTED FROM PREDICTORS =        .333
REDUNDANCY OF PREDICTORS GIVEN CRITERIA        .205
NOTE THAT ALL VALUES ARE PROPORTIONS
FACTOR STRUCTURE FOR CRITERIA, COLUMNS ARE FACTORS
RDIVLOSSTOGR         .955      .296
RDIVLOSS2AIR       -1.000      .020
FACTOR   VARIANCE EXTRACTED    REDUNDANCY
   1               .956            .945
   2               .044            .011
TOTAL VARIANCE EXTRACTED FROM CRITERIA =      1.000
TOTAL REDUNDANCY FOR CRITERIA GIVEN PREDICTORS        .956
NOTE THAT ALL VALUES ARE PROPORTIONS
WILKS LAMBDA FOR TOTAL SET =         .0085987
CHI SQUARE FOR TOTAL =    359.0893059
N.D.F. =     12
```

Table XXIX, Observed and Predicted Values for Output Variables
(continued on next page)

```
OBS RDIVLOSSTOGR
OBS RDIVLOSS2AIR
PRD RDIVLOSSTOGR
PRD RDIVLOSS2AIR
(10X,5E14.7/(10X,5E14.7))
```

.1138006+02	.4940103+01	.1269389+02	.5260203+01
.1125998+02	.3990103+01	.1165983+02	.4131180+01
.9740010+01	.6850084+01	.1004735+02	.6966059+01
.1367996+02	.1339982+01	.1461232+02	.5831017+00
.1075994+02	.6629987+01	.1109419+02	.6414734+01
.8929920+01	.9600016+01	.8694559+01	.9359606+01
.1476003+02	.1749976+01	.1405260+02	.1613172+01
.1050008+02	.8460119+01	.9742887+01	.8493356+01
.1032998+02	.6549882+01	.1060559+02	.6250914+01
.1288992+02	.2820006+01	.1378137+02	.3136812+01
.9769988+01	.7240116+01	.1020861+02	.7289008+01
.1040997+02	.6050057+01	.1041109+02	.5931622+01
.1189996+02	.5059876+01	.1243006+02	.5215433+01
.1229995+02	.3809930+01	.1211964+02	.4180316+01
.1070992+02	.6980094+01	.9851367+01	.6961692+01
.1455001+02	.1570060+01	.1434701+02	.8532572+00
.1165997+02	.4550070+01	.1156140+02	.4895460+01
.9009918+01	.1056000+02	.8492657+01	.1060872+02
.1478007+02	.2000017+01	.1414511+02	.2102375+01
.1109998+02	.4999990+01	.1129609+02	.5165616+01
.9379931+01	.9619978+01	.8959870+01	.9089450+01
.1417006+02	.2650071+01	.1358687+02	.2817520+01
.9749945+01	.8849895+01	.9937384+01	.8812648+01
.1170005+02	.4440022+01	.1087681+02	.4727274+01
.1179992+02	.5179906+01	.1223408+02	.5211067+01
.1149997+02	.3939941+01	.1185582+02	.4135546+01
.1021007+02	.7290022+01	.1031117+02	.7010829+01
.1465005+02	.2159971+01	.1405240+02	.2545447+01
.9600057+01	.8110011+01	.9908687+01	.8162172+01
.1115993+02	.5650043+01	.1043999+02	.5649822+01
.1236008+02	.4290049+01	.1288647+02	.4607894+01
.1096996+02	.5090076+01	.1158798+02	.5073700+01
.1056997+02	.6580081+01	.9926617+01	.6675848+01
.1386000+02	.1300057+01	.1457168+02	.2659373+00
.1162999+02	.5679987+01	.1107610+02	.5784800+01
.8860030+01	.1051009+02	.8753285+01	.1030670+02
.1425006+02	.1789901+01	.1417400+02	.1820575+01
.1155993+02	.4270087+01	.1156712+02	.4574252+01
.9319976+01	.9809876+01	.8659945+01	.9962615+01
.1346001+02	.2380068+01	.1366212+02	.2531675+01
.9810074+01	.8259984+01	.1039379+02	.8205108+01
.1080003+02	.5670006+01	.1034516+02	.5620658+01
.1117004+02	.4810092+01	.1249471+02	.4909048+01
.1279998+02	.3190076+01	.1208049+02	.3548226+01
.8709968+01	.8070086+01	.9825872+01	.7900168+01
.1191007+02	.5920046+01	.1239396+02	.6133368+01

Table XXIX, Observed and Predicted Values for Output Variables (concluded)

.1189996+02	.3649976+01	.1168873+02	.3849380+01
.1053999+02	.6300097+01	.1031838+02	.6374694+01
.1407996+02	.1789901+01	.1408066+02	.1476485+01
.1138006+02	.6139887+01	.1116944+02	.6128889+01
.9459929+01	.9410118+01	.9150964+01	.8752067+01
.1365992+02	.2459917+01	.1356729+02	.2502511+01
.9609991+01	.8070086+01	.1000351+02	.8191336+01
.1112996+02	.5689968+01	.1083026+02	.5663594+01
.1465005+02	.1440049+01	.1473156+02	.1188238+01
.1094992+02	.5390022+01	.1090902+02	.5498633+01
.8730011+01	.9710065+01	.8760489+01	.9670570+01
.1434993+02	.1850044+01	.1398795+02	.1919557+01
.9699924+01	.9050030+01	.9782036+01	.9125446+01
.1038000+02	.5210105+01	.1063108+02	.5312438+01
.1213996+02	.4130095+01	.1286689+02	.4292885+01
.1186998+02	.4329974+01	.1165411+02	.4452389+01
.9440061+01	.7780121+01	.9880064+01	.7612168+01
.1192994+02	.5640062+01	.1249451+02	.5841323+01
.1369007+02	.2940036+01	.1224629+02	.3217043+01
.1008005+02	.7309984+01	.9660272+01	.7299076+01
.1405991+02	.1629947+01	.1436658+02	.1168266+01
.1093005+02	.5350097+01	.1149527+02	.5516772+01
.8769923+01	.9720046+01	.8539211+01	.9672404+01
.1397991+02	.2209877+01	.1362149+02	.2214511+01
.1001992+02	.7339927+01	.1037570+02	.7575174+01
.1038993+02	.6890008+01	.1040388+02	.6567757+01
.1367003+02	.3149895+01	.1339429+02	.3469829+01
.9619926+01	.7879933+01	.1000923+02	.7870127+01
.1266003+02	.4789874+01	.1099755+02	.5017485+01
.1184001+02	.5350097+01	.1227472+02	.5528231+01
.1167008+02	.4609957+01	.1187391+02	.4765480+01
.1062992+02	.5900084+01	.1025245+02	.6063730+01
.1405991+02	.1710052+01	.1414531+02	.1170100+01
.1157003+02	.5350097+01	.1113029+02	.5496799+01
.8910051+01	.1022013+02	.9125469+01	.9690542+01

Further, Figures XVIII-XXI graphically display the values of the observed and predicted variables, while Figures XXII and XXIII are graphs of predicted values versus observed values for the variables RDIVLOSSTOGR and RDIVLOSSTOAI. Additionally, these latter figures contain the simple linear regression line, its coefficients, and the correlation coefficient between the observed and predicted values. Figures XVIII-XXIII were generated by MIDAS, an interactive graphics capability available at CAA. MIDAS, TAGS and MAS were integrated in the purpose of creating these plots.

7. Applicability of the Multivariate Analysis System. The Multivariate Analysis System is designed to aid the analyst in disentangling the complex interrelationships inherent in a multivariate data base. In the case at hand, the data base consists of 81 observations on each of 19 variables, 6 of which are input variables and 13 output variables. An examination of this data base reveals its complexity and the opportunity to analyze these complexities. This problem is amenable to techniques which form the Multivariate Analysis System.

8. Observations on Correlation Coefficients. Tables VI and VII display different views of the matrix of intervariable correlation coefficients. Table VI displays the complete matrix in upper triangular form, while Table VII displays only those correlation coefficients which exceed a user-supplied threshold.

a. Threshold. In this work, the threshold was chosen to by 0.80. This value represents a level, chosen subjectively, above which two variables are considered to carry essentially the same information.

b. Data Reduction. Twenty (or 6.9 percent) of the 289 correlation coefficients, exceed the stated threshold. One common practice in the art of data reduction is to eliminate from the data base variables having a high correlation with several other variables. With this in mind, the original variables BTOTGDLOSS RTOTGDLOSS were deleted from consideration.

c. Losses. There is a linear connection between Red's initial aircraft inventory and Red's bomber losses. Red's losses of bombers are a linear function of Red's initial aircraft inventory, a not wholly unexpected result.

d. Concomitants. FEBA movement has a high negative correlation with both the number of Blue divisions lost to ground forces and the number of Blue divisions lost to air forces. As the level of Blue's divisional losses rises, due to ground or air causes, FEBA progressively tends to more negative numbers. This, too, is an expected result. Note, however, the large negative correlation between FEBA and Red's division loss to ground forces.

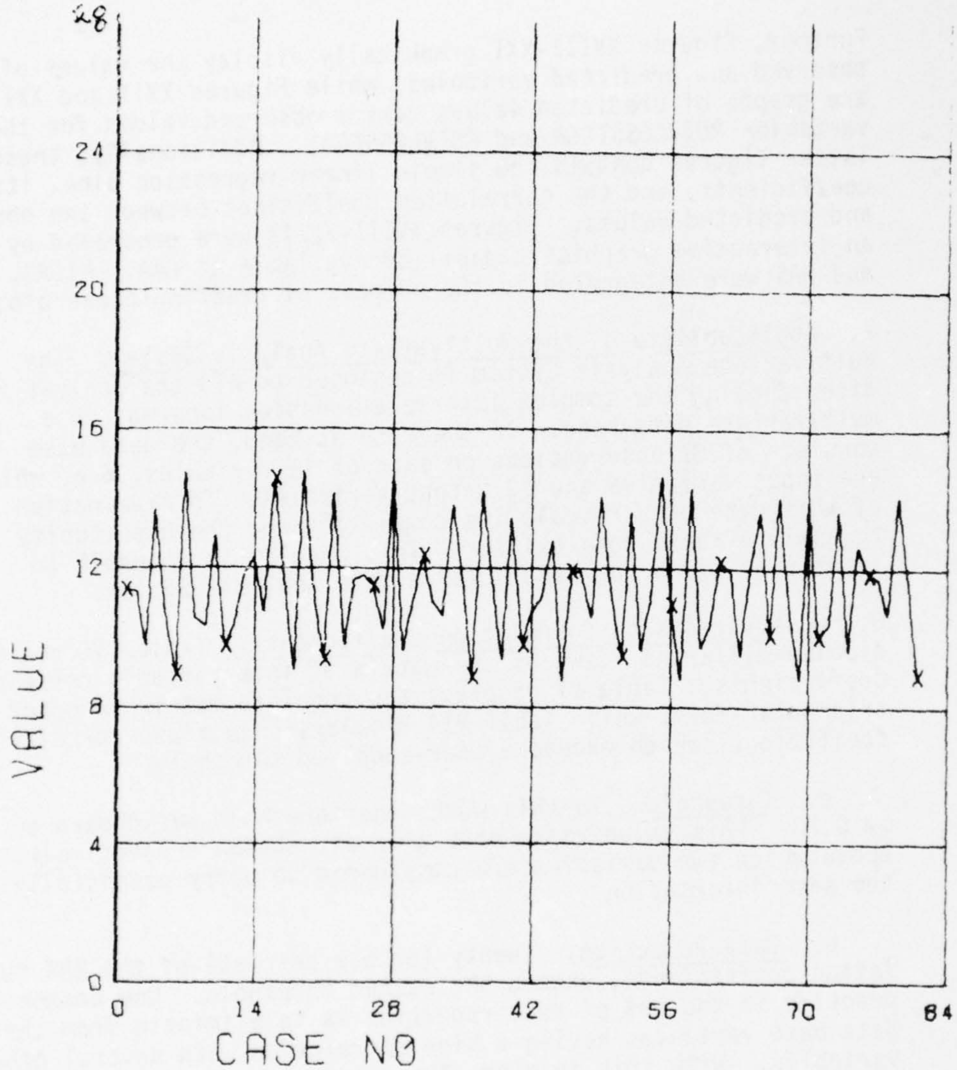


Figure XVIII, Observed RDIVLOSSTOGR

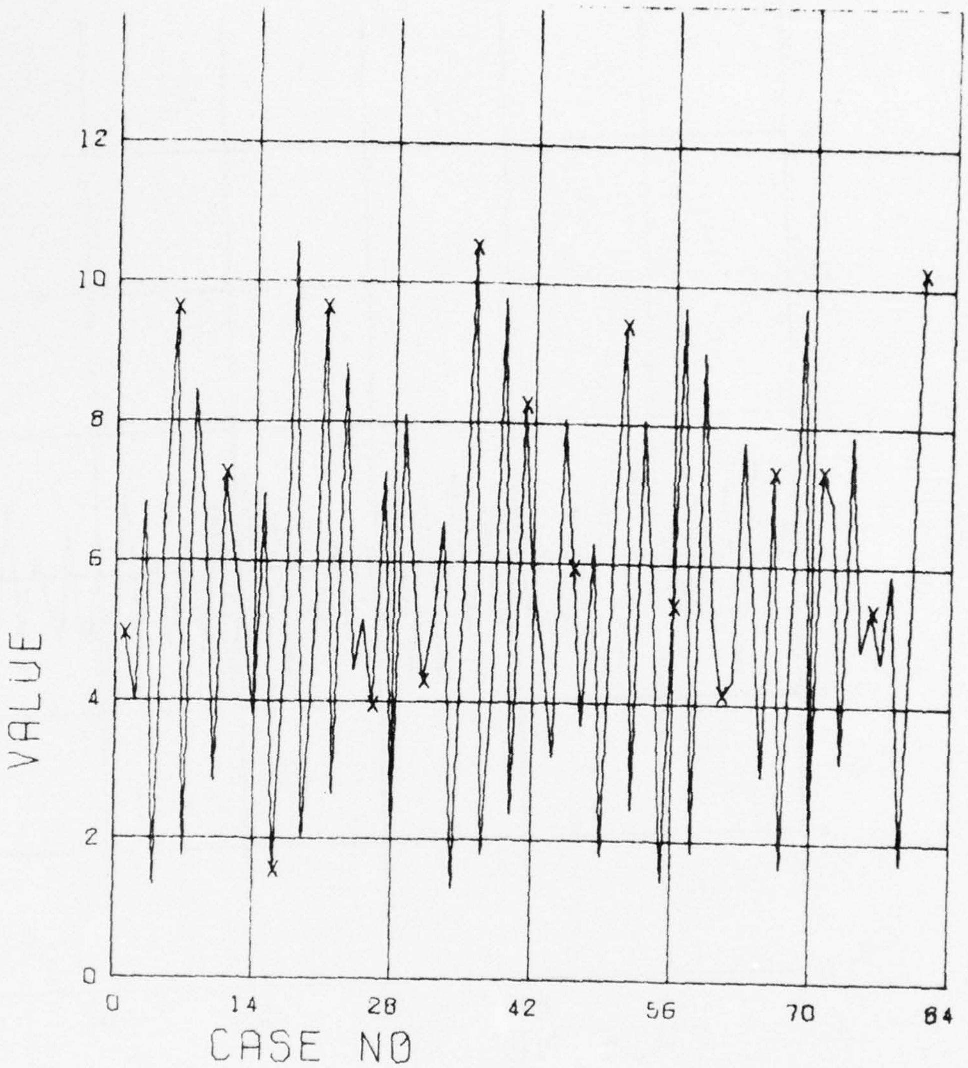


Figure XIX, Observed RDIVLOSS2AIR

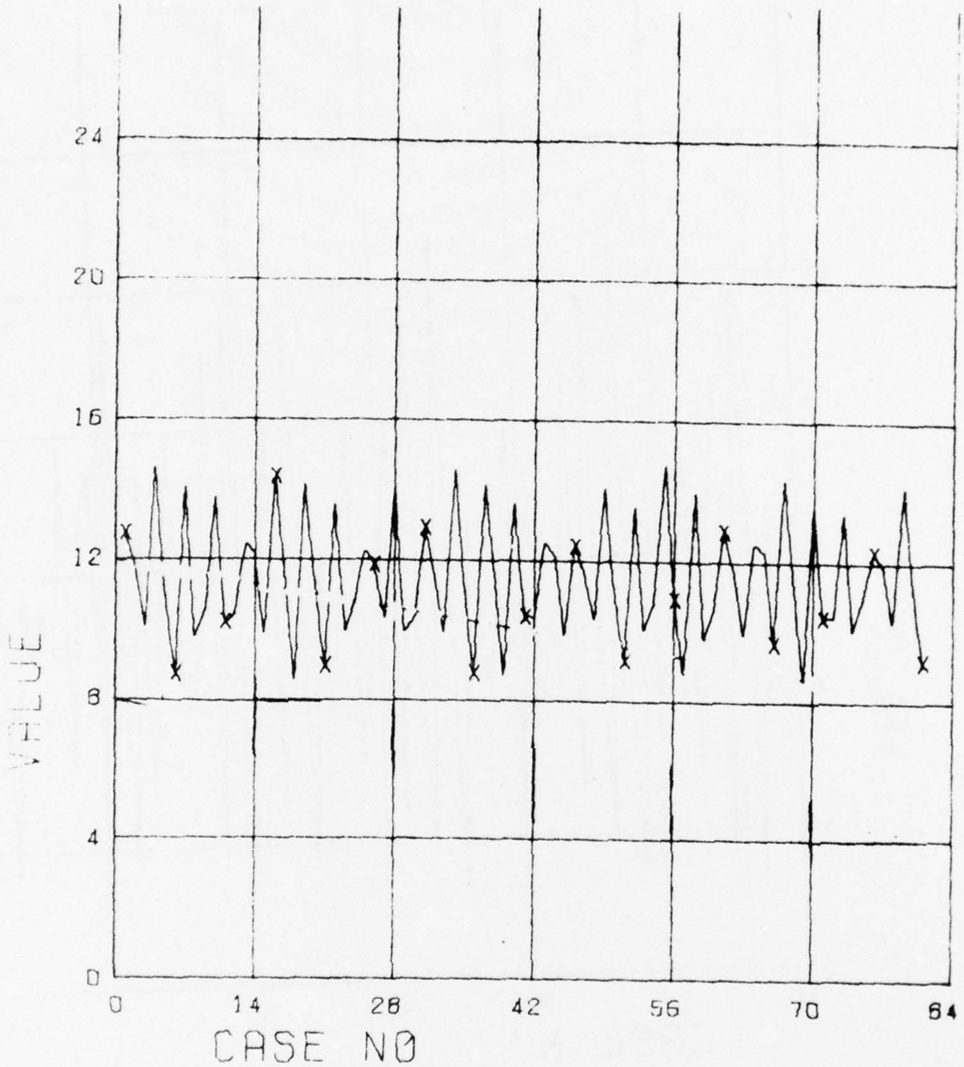


Figure XX, Predicted RDIVLOSSTOGR

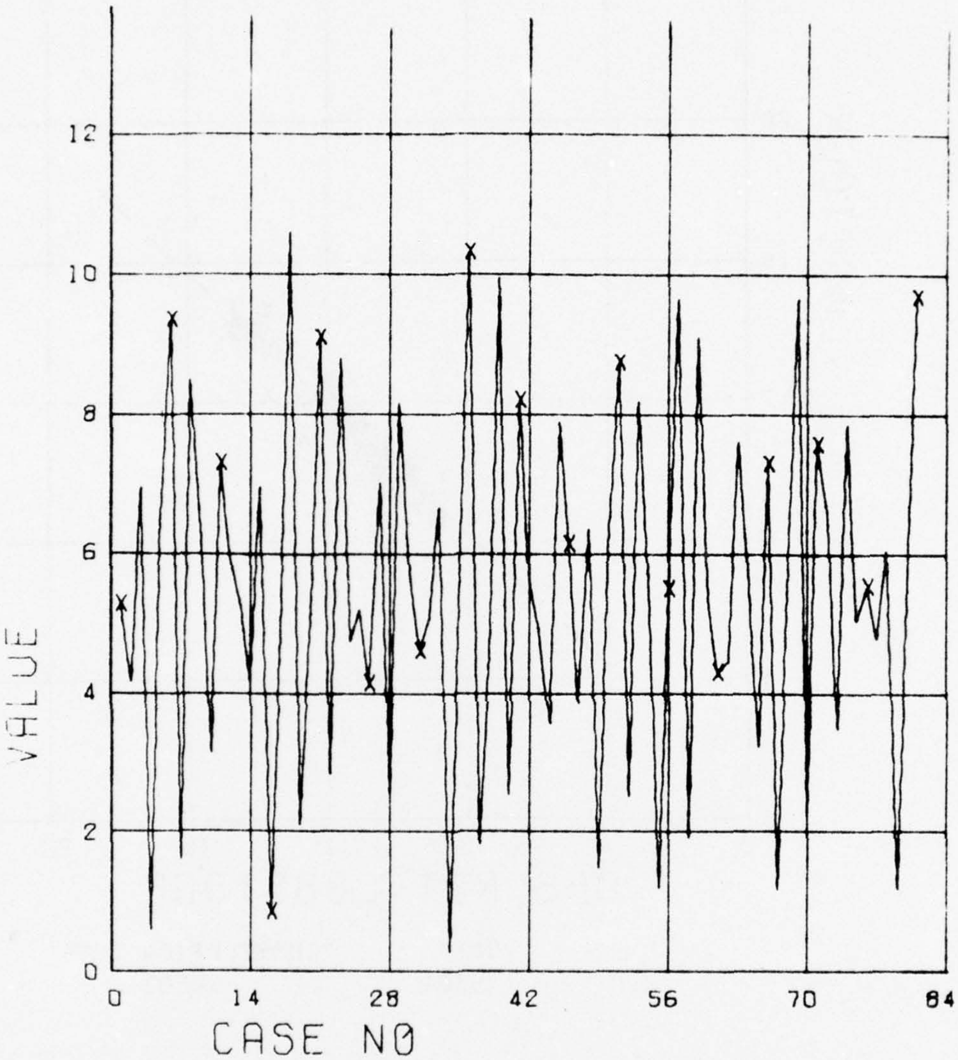
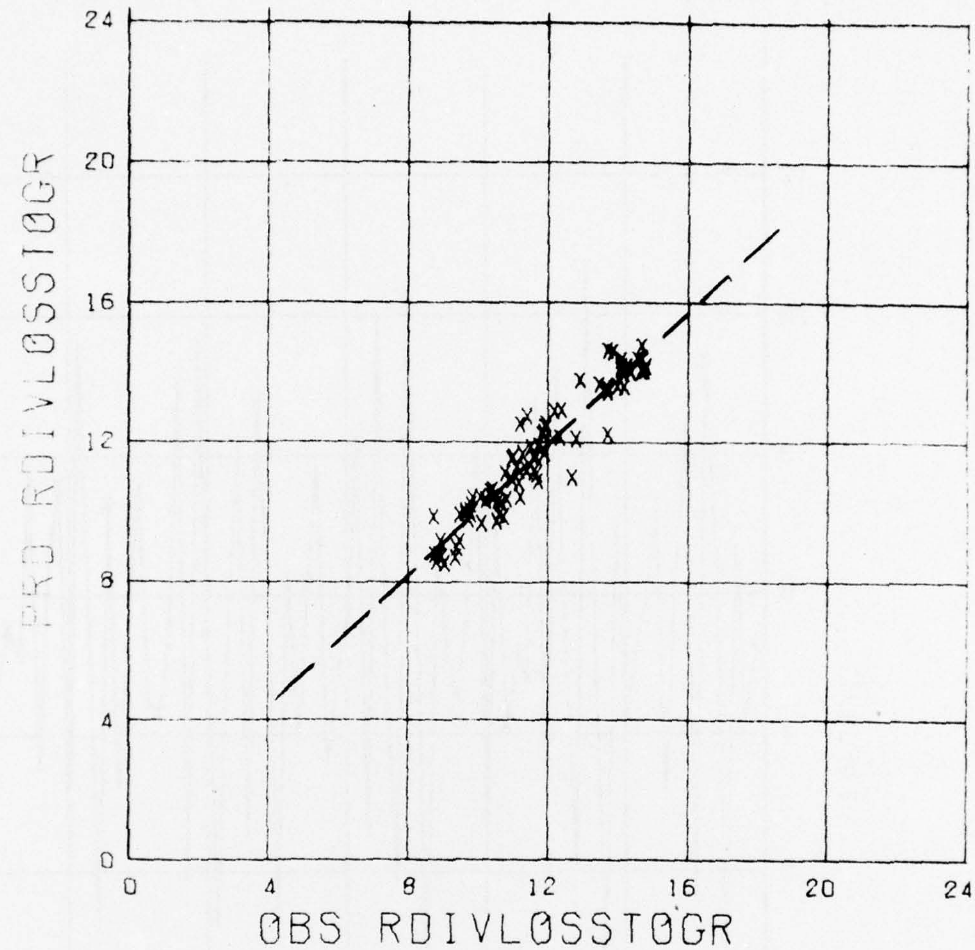


Figure XXI, Predicted RDIVLOSS2AIR



A	B(X)	CORRELATION COEF
.57	.9501	.9501

Figure XXII, Plot of Observed vs Predicted Values for RDIVLOSSTOGR With Simple Linear Regression

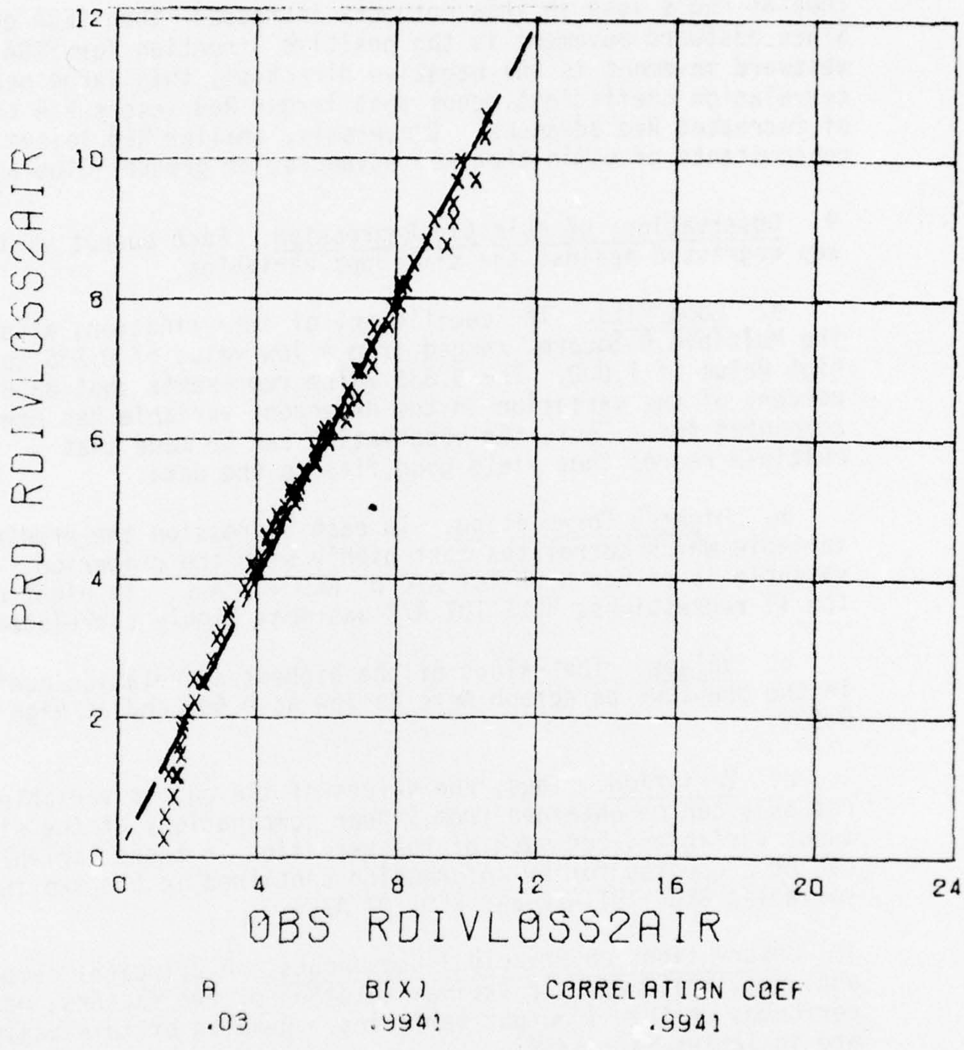


Figure XXIII, Plot of Observed vs Predicted Values for RDIVLOSS2AIR With Simple Linear Regression

Thus as Red's loss in this category increases, then FEBA decreases. Since eastward movement is the positive direction for FEBA and westward movement is the negative direction, this large negative correlation coefficient means that larger Red losses are concomitants of increased Red advances. Conversely, smaller Red losses are concomitants of diminished Red advances, or greater Blue advances.

9. Observations on Multiple Regression. Each output variable was regressed against the six input variables.

a. Good Fits. The coefficient of determination, also called the Multiple R-Square, ranged from a low value of 0.836 to a high value of 1.000. The 0.836 value represents that 83.6 percent of the variation in the dependent variable has been accounted for. Thus, the observation can be made that multiple regressions yield good fits to the data.

b. Highest Correlation. In each regression the predictor variable which correlates most highly with the criterion variable is either BLUE TOT A/C or RED TOT A/C. In eight of the 11 regressions, BLUE TOT A/C was most highly correlated.

c. Values. The values of the highest correlation coefficients in the previous paragraph were as low as 0.695 and as high as 0.999.

d. Variation. Thus, the values of the output variables not only can be obtained from linear combinations of the six input variables, but much of the variation in these variables can be accounted for by information contained in the two input variables BLUE TOT A/C and RED TOT A/C.

10. Observations on Principal Components. A principal components analysis, followed by a varimax rotation of two factors, was performed on the 11 output variables. Results of this analysis are in Tables XX - XXIV.

a. Eigenvalues. Since there are 11 variables, there are 11 principal components. Each principal component is associated with exactly one of the eigenvalues of the correlation matrix.

b. Relationships. Table XXIII displays the correlation coefficients between each principal component and underlying variable. In this table the columns represent principal components while the rows represent underlying variables. For example, the first principal component correlates 0.883 with FEBA, 0.713 with BFTRLOSS, and 0.319 with BATTLOSS.

Thus, using 0.75 as a threshold above which a correlation coefficient is considered high, Table XXIII shows that principal component number one has high correlation coefficients with FEBA, BDIVLOSSTOGR, RFTRLOSS, RDIVLOSSTOGR, and RDIVLOSS2AI. By the same standard, the second principal component correlates highly with BATTLOSS and RBMRLOSS.

c. Factor Pattern Matrix. An examination of the entire factor pattern matrix in Table XXIII, as in above paragraph, shows the degree to which each principal component is correlated with the original output variables. The nonzero entries in the matrix clearly show that each output variable is represented in the complex structure of each principal component. However, Table XXIII also shows that no principal component after the second one shares a correlation coefficient with any output variable higher than 0.75.

d. Magnitude. Table XXII shows that the first two principal components account for 94 percent of the variation of the original variables. A standard practice in factor analysis is to discard principal components associated with eigenvalues of less than 1.00 magnitude. The decision was made to retain only the first two principal components for purposes of further analysis.

e. Loadings. The purpose of a varimax rotation is to maximize the variance of squared loadings for each factor. A new set of factors then preserves the information carried by the principal components for which the factor pattern has very high and very low loadings. Improved interpretability for the factors results.

f. Rotation. The results of a varimax rotation on two principal components are shown in Table XXIV. It can be seen from the new factor pattern that the first factor correlates highly with the five output variables, BDIVLOSS2AI, RDIVLOSSTOGR, BDIVLOSSTOGR, FEBA, and RDIVLOSS2AIR. These variables are descriptors of the ground battle. Similarly, the second factor correlates highly with the output variables RBMRLOSS, RFTRLOSS, BATTLOSS, BFTRLOSS, RATTLOSS, and BBMRLOSS. These variables are descriptors of the air battle. These observations suggest that the first factor be named "Ground Losses" and the second factor be named "Aircraft Losses."

g. Losses. In summary, 94 percent of the variation in the values of the original 11 output variables can be accounted for by two factors, Ground Losses and Aircraft Losses.

11. Observations on Canonical Correlation. Four distinct canonical correlations were performed. Only the fourth case is presented in detail.

a. Fourth Case. The set of independent, or predictor, variables for the fourth case consists of BDIV, BAC, BSHLT, RDIV, RAC, and RSHLT. The set of dependent, or criterion, variables for the fourth case consists of RDIVLOSSTOGR and RDIVLOSS2AIR.

b. Maximization. The canonical correlation program produced a linear combination of the six predictor variables, and a linear combination of the two criterion variables, so that these two linear combinations have a maximized correlation coefficient between them. Table XXVIII shows that the results of this computation are that (0.004) BLUE DIV + (-0.737) BLUE TOT A/C + (-0.146) BL A/C SHELT + (0.004) RED DIV + (0.645) RED TOT A/C + (0.144) RED A/C SHLT and (0.063) (RDIVLOSSTOGR) + (-0.940) RDIVLOSS2AIR have a correlation coefficient of 0.99432. Similarly, the next best pair of canonical variates is given by the second column of weights in Table XXVIII, and these variates have a canonical correlation coefficient equal to 0.49117.

c. Variates. The first canonical variates, with its high canonical correlation coefficient, establish a strong link between the dimensions of the problem (1) inherent in the two criterion variables, and (2) defined by the inputs of the predictor variables. These dimensions can be labeled, respectively, the "loss structure" and the "input force structure."

d. Predicted Values. Using the structure produced by the canonical correlation program, it is possible to generate predicted values for RDIVLOSSTOGR and RDIVLOSS2AIR from any observed values of the input variables. Table XXIX displays the observed values for these variables alongside the predicted values for the 81 cases in the data base. Columns 1 and 2 are the observed values for RDIVLOSSTOGR and the observed values for RDIVLOSS2AIR, respectively. Columns 3 and 4 are the predicted values for these variables, respectively. A comparison of Columns 1 and 3 reveals a good fit of predicted values of RDIVLOSSTOGR to observed values of RDIVLOSSTOGR. Similarly, a comparison of Columns 2 and 4 reveals a close relationship between the observed values of RDIVLOSS2AIR and the predicted values for this variable.

e. Graphs. In order to gain further insight into the data of Table XXIV, graphs were plotted which display the case-by-case variation for each of the four variables under discussion. Figures XVIII and XX represent the graphs for the observed and predicted values, respectively, for RDIVLOSSTOGR.

f. Similarity. As these graphs show, the observed and predicted values not only fall in similar ranges, but they increase and decrease together. Therefore, a high correlation coefficient exists between these two variables. This issue will be pursued below. In the same way, there appears to be great similarity in the behaviors exhibited in Figures XIX and XXI.

g. Behavior. A high degree of similarity of behavior appears to exist between the observed and predicted values of RDIVLOSSTOGR. In order to display this relationship between these two variables, a graph of observed RDIVLOSSTOGR versus predicted RDIVLOSSTOGR was plotted. In Figure XXII, the observed values of RDIVLOSSTOGR were plotted as the abscissa, and the predicted values of RDIVLOSSTOGR were plotted as the ordinate. The points appear to be located in such a way that a straight-line fit would be appropriate. Indeed, this was done and the results of this simple linear regression are also displayed in Figure XXII.

h. Summary of Fit. The straight-line fit, which was introduced in paragraph g above, may be summarized as follows:

Predicted RDIVLOSSTOGR = 0.57 + (0.9501) observed RDIVLOSSTOGR. Additionally, the correlation coefficient between the observed and predicted variables equals 0.9501.

APPENDIX A

STUDY CONTRIBUTORS

AN APPLICATION OF THE MULTIVARIATE
ANALYSIS SYSTEM TO TAGS

APPENDIX A
STUDY CONTRIBUTORS

1. Study Director

Dr. Daniel A. Nussbaum, Methodology, Resources and Computation Directorate

2. Study Team

Ms. Ruth W. Meads, Project Planning and Control Office

3. Support Personnel

Ms. Darrie-Ann Anderson

Ms. Linda L. Prieto

Ms. Thelma Laufer

Ms. Wanda Lynch

APPENDIX B

REFERENCES

AN APPLICATION OF THE MULTIVARIATE ANALYSIS SYSTEM TO TAGS

APPENDIX B

REFERENCES

1. Emerson, D. E., "TAGS-V: A Tactical Air-Ground Warfare Model," RAND Report R-1242-PR, RAND Corporation, June 1973.

2. Nussbaum, D. A., "Multivariate Analysis System," CAA-D-76-4, US Army Concepts Analysis Agency, Bethesda, MD., June 1976.

3. "Fractional Factorial Experiment Designs for Factors a Three Levels," US Department of Commerce, National Bureau of Standards, Applied Mathematics Series No. 54.

APPENDIX C

DISTRIBUTION

AN APPLICATION OF THE MULTIVARIATE
ANALYSIS SYSTEM TO TAGS

APPENDIX C
DISTRIBUTION

Addressee	Copies
Defense Documentation Center for Scientific and Technical Information Building 5 Cameron Station Alexandria, VA 22314	12
The Army Central Library (ASDIRS) Room 1A600 Pentagon Building Washington, DC 20310	1
Deputy Under Secretary of the Army (Operations Research) HQ DA, Room 2E621 Pentagon Building Washington, DC 20310	1
Deputy Chief of Staff for Operations and Plans (DAMO-ZD) Room 3A538 Pentagon Building Washington, DC 20310	1
Director US Army TRADOC Systems Analysis Activity White Sands Missile Range New Mexico 88002	1
US Army Concepts Analysis Agency	
Chief, AD, ATTN: Library Branch	1
Director, FD	3
Director, JF	1
Director, SM	1
Director, WG	1
Director, MR	15